

Kognitive Täuschungen vor Gericht

Eine empirische Studie

286 + XL Seiten, 24 Abbildungen, 25 Tabellen

Dissertation

der Rechtswissenschaftlichen Fakultät der Universität

Zürich zur Erlangung der Würde eines Doktors der

Rechtswissenschaft

vorgelegt von

Mark Daniel Schweizer

von Untereggen (SG)

Promotionsdatum: 29. Juni 2005

Dank

Eine Arbeit wie die vorliegende wäre nicht denkbar, wenn sich nicht Richterinnen und Richter fänden, die sich die Zeit nehmen, einen Fragebogen auszufüllen. Mein Dank geht daher zuerst und vor allem an alle 415 Richterinnen und Richter der Kantone Aargau, Bern, beider Basel, Graubünden, St. Gallen und Zürich, die sich an einer der beiden Umfragen im Herbst 2003 und Herbst 2004 beteiligt haben.

Weiter bedanke ich mich bei Prof. Manfred Rehbinder und dem Europäischen Institut für Rechtspsychologie, die diese Arbeit von Anfang an unterstützt haben, Dr. Lukas Furrer für die Hilfe bei der statistischen Auswertung der Umfrage, lic. phil. Ruth Osterwalder für das sorgfältige Lektorat, Barbara Marti und Martina Hrabovecka für die Hilfe beim Beschaffen der Literatur und meinen ehemaligen Arbeitskollegen bei Meyer Lustenberger Rechtsanwälte, Zollikon, für die Teilnahme an den Pretests.

Inhaltsübersicht

Inhaltsverzeichnis	II
Literaturverzeichnis	X
Quellenverzeichnis	XXXIX
Allgemeiner Teil	1
§ 1 Einleitung, Aufbau der Arbeit und Eingrenzung des Themas	1
§ 2 American Legal Realism	5
§ 3 Ökonomische Analyse des Rechts	10
§ 4 Rezeption der psychologischen Forschung durch die Rechtswissenschaft: <i>Behavioral Law and Economics</i> oder <i>Empirical Legal Realism</i>	35
§ 5 Kritik	41
§ 6 Prospect Theory	53
Besonderer Teil: Einzelne kognitive Täuschungen und ihr Einfluss auf das Urteil von Richtern und Parteien in Gerichtsverfahren	64
§ 7 Methode, Auswertung, interne und externe Validität der Studie	64
§ 8 Ankereffekt (<i>anchoring</i>)	77
§ 9 Darstellungseffekt (<i>framing</i>)	90
§ 10 Unterlassungseffekt (<i>omission bias</i>)	108
§ 11 Denken in Repräsentativität	124
§ 12 Bestätigungsfehler (<i>confirmation bias</i>)	178
§ 13 Nichtbeachten der Regression zum Mittelwert	196
§ 14 Rückschaufehler (<i>hindsight bias</i>)	209
§ 15 Hofeffekt (<i>halo effect</i>)	230
§ 16 Kontrast- und Kompromisseffekt	243
§ 17 Selbstüberschätzung und <i>overconfidence</i>	261
§ 18 Zusammenfassung/Thesen	271
Anhang A: Herleitung des Bayes-Theorems aus den Axiomen der Wahrscheinlich- keitstheorie	275
Anhang B: Fragebogen 2003	278
Anhang C: Fragebogen 2004	282
Lebenslauf	286

Inhaltsverzeichnis

Literaturverzeichnis	X
Quellenverzeichnis.....	XXXIX
Allgemeiner Teil.....	1
§ 1 Einleitung, Aufbau der Arbeit und Eingrenzung des Themas	1
I. Einleitung.....	1
II. Aufbau der Arbeit und Eingrenzung des Themas	3
§ 2 American Legal Realism	5
§ 3 Ökonomische Analyse des Rechts	10
I. Ursprünge der Ökonomischen Analyse des Rechts	10
II. Erfolg der Ökonomischen Analyse des Rechts in den USA.....	13
III. Gründe für den Erfolg der Ökonomischen Analyse des Rechts	16
IV. Rationales Verhalten im Sinne der ökonomischen Theorie.....	17
A. Definitorisches Modell.....	18
B. Erwartungsnutzenmodell (<i>expected utility theory</i>).....	18
C. Eigeninteresse.....	21
D. Vermögensmaximierung	21
V. Kritik an der Ökonomischen Analyse des Rechts	22
A. Kritik an der Effizienz als Rechtsziel	22
B. Kritik am Homo Oeconomicus.....	23
1. Kritik an der Annahme (ausschliesslich) eigeninteressierten Verhaltens.....	24
2. Kritik an der Annahme (ausschliesslich) rationalen Verhaltens	25
a. Eingeschränkte Rationalität (<i>bounded rationality</i>).....	25
b. Heuristics and biases	27
c. Referenzpunkte, <i>Framing</i> und Risikoverhalten.....	30
d. Kontext und Präferenzen	31
e. Weitere „Kognitive Täuschungen“	32
f. Zusammenfassung.....	33
§ 4 Rezeption der psychologischen Forschung durch die Rechtswissenschaft: <i>Behavioral Law and Economics</i> oder <i>Empirical Legal Realism</i>	35
I. Behavioral Law and Economics.....	35

II. Empirical Legal Realism?	39
§ 5 Kritik	41
I. Kritik an der psychologischen Forschung zu den <i>heuristics and biases</i>	41
A. Zu pessimistische Sicht	41
B. Der „richtige“ normative Standard	42
C. Generalisierbarkeit	42
D. Theoriedefizit	43
E. Duales System.....	44
II. Kritik an <i>Behavioral Law and Economics</i>	47
III. Stellungnahme	51
§ 6 Prospect Theory – ein deskriptives Modell menschlichen Risikoverhaltens	53
I. Ein Gedankenexperiment zum Einstieg	53
A. Das Allais Paradox	55
II. Prospect Theory	56
A. Referenzpunkt	56
B. Wertfunktion.....	57
1. Loss aversion	58
C. Gewichtungsfunktion	60
D. Kombination von Wert- und Gewichtungsfunktion.....	61
E. Die Unterscheidung zwischen Risiko und Unsicherheit	62
F. Anwendungsbeispiele.....	63
Besonderer Teil: Einzelne kognitive Täuschungen und ihr Einfluss auf das Urteil von Richtern und Parteien in Gerichtsverfahren	64
§ 7 Methode, Auswertung, interne und externe Validität der Studie	64
I. Methode.....	64
II. Auswertung, oder: was bedeutet „statistisch signifikant“?	66
III. Interne Validität	69
IV. Externe Validität	70
§ 8 Ankereffekt (<i>anchoring</i>)	77
I. Ankereffekt vor Gericht	79
II. Eigene Studie zum Ankereffekt.....	82
A. Frage zum Ankereffekt.....	82

B. Resultate.....	83
C. Diskussion.....	85
1. Wie das kontradiktorische Verfahren den Einfluss des Ankereffekts ver- ringern kann.....	85
2. Rechtliche Rahmenbedingungen, die den Einfluss des Ankereffekts mässigen.....	85
III. Ankereffekt in Verhandlungen.....	87
1. Tipps für den gewinnbringenden Einsatz des Ankereffekts in Verhandlungs- situationen	88
a. Nach den Sternen greifen	88
b. Das erste Angebot machen	88
c. Nicht zu bescheiden sein	88
2. Was man tun kann, um den Einfluss des Ankereffekts zu verringern.....	89
§ 9 Darstellungseffekt (<i>framing</i>).....	90
I. Der Darstellungseffekt im Recht.....	90
A. Laborexperimente	90
B. Feldstudien	91
II. Unterliegen auch Richter dem Darstellungseffekt?.....	93
A. Amerikanische Studie von Guthrie und Kollegen.....	94
III. Eigene Studie zum Darstellungseffekt.....	95
A. Frage zum Darstellungseffekt	95
B. Resultate.....	96
C. Diskussion	97
IV. Prospect Theory und Klagen mit geringen Erfolgsaussichten.....	97
V. Prospect Theory und Steuerhinterziehung.....	99
VI. Status quo Bias und Vertragsverhandlungen	101
A. Demonstration des Status quo Biases bei Präferenzen für Vertragsklauseln.....	102
B. Die Bevorzugung des Status quo als Folge der Abneigung gegen Handlung.....	102
C. Dispositives Recht, Allgemeine Vertragsbedingungen und Status quo	104
1. Empirische Studien von Russell Korobkin	104
2. Schlussfolgerung für den Praktiker.....	106
§ 10 Unterlassungseffekt (<i>omission bias</i>).....	108

I.	Psychologische Studien zur moralischen Bewertung von Handlungen und vorsätzlichen Unterlassungen.....	108
II.	Rechtliche Beurteilung von Handlungen und vorsätzlichen Unterlassungen	111
	A. Unterlassung der Nothilfe (Art. 128 StGB).....	112
	B. Unechte Unterlassungsdelikte.....	113
III.	Eigene Studie zum Unterlassungseffekt.....	115
	A. Frage zum Unterlassungseffekt	115
	B. Rechtliche Beurteilung der Sachverhaltsvarianten	116
	C. Resultate	117
	D. Diskussion	120
	1. Unterlassungseffekt und Sterbehilfe	121
§ 11	Denken in Repräsentativität	124
I.	Ein Beispiel zum Einstieg	124
II.	Der normative Standard: das Bayes-Theorem.....	125
	A. Die Lösung des Taxi-Problems nach dem Bayes-Theorem	127
III.	Das deskriptive Modell: Denken in Repräsentativität	128
	A. Typische Fehler, die aus dem repräsentativen Denken resultieren.....	131
	1. Ausländer sind kriminell, Marihuana ist eine Einstiegsdroge und Videospiele machen gewalttätig.....	131
	2. Fallacia Consequentis.....	132
	3. Was typisch ist, ist beweiskräftig.....	133
	4. Der Trugschluss des Anklägers	135
	5. Umgekehrter Fehler: was untypisch ist, ist unwahrscheinlich	137
	B. Kritik am deskriptiven Modell in der Psychologie	139
	1. Was heisst „ein Ereignis tritt mit einer Wahrscheinlichkeit von 0,x auf“?	140
	a. Objektive Wahrscheinlichkeit.....	140
	b. Subjektive Wahrscheinlichkeit	141
	c. Wahrscheinlichkeitsbegriff und Tatsachenfeststellung vor Gericht	142
	2. Häufigkeiten statt Wahrscheinlichkeiten.....	146
IV.	Das Bayes-Theorem und richterliche Beweiswürdigung.....	148
	A. US-amerikanische Lehre	148
	B. Schweizer Lehre.....	150

C. Deutsche Lehre	150
D. Stellungnahme	151
E. Das Problem der a priori Wahrscheinlichkeiten.....	152
1. Die massgebliche Referenzklasse.....	152
2. Geschätzte Anfangswahrscheinlichkeiten, oder „garbage in, garbage out“?. 153	
F. Dogmatische Erkenntnisse	158
V. Eigene Studie zur Repräsentativitätsheuristik.....	161
A. Frage zur Repräsentativitätsheuristik.....	161
B. Resultate.....	162
C. Diskussion	162
VI. Wie kann man Fehler des repräsentativen Denkens vermeiden?.....	163
A. Gutachter sollten nie die Belastungswahrscheinlichkeit angeben.....	163
B. Statistische Informationen sollten in natürlichen Häufigkeiten präsentiert werden.....	167
VII. Beweismwürdigung und DNA-Analysen.....	169
A. Die Technik der DNA-Analyse	169
B. Der Beweiswert von DNA-Analysen und falsch positive Resultate	170
§ 12 Bestätigungsfehler (<i>confirmation bias</i>)	178
I. Informationssuche	179
A. Pseudodiagnostizität	179
B. Bevorzugte Suche nach bestätigenden Informationen.....	181
II. Informationsbewertung.....	185
A. Ambivalente Informationen werden als Bestätigung interpretiert	185
B. Bestätigende Informationen werden stärker gewichtet als widersprechende Informationen	186
III. Folgen der Bestätigungstendenzen.....	188
A. Festhalten selbst an einer diskreditierten Hypothese (<i>belief perseverance</i>)	188
B. Einfluss der Reihenfolge der Beweismittel (<i>primacy effect</i>)	188
C. Zu grosse Sicherheit im Urteil (<i>overconfidence</i>).....	190
IV. Eigene Studie zum Bestätigungsfehler	190
A. Frage zur Pseudodiagnostizität.....	190
B. Resultate.....	192

C. Diskussion	193
V. Was kann gegen Bestätigungstendenzen unternommen werden?	194
A. Kontrafaktisches Denken (<i>counterfactual reasoning</i>)	194
B. Alternative Erklärungen generieren und testen	194
§ 13 Nichtbeachten der Regression zum Mittelwert	196
I. Regression zum Mittelwert und harte Strafen	199
A. Das Problem der Raserunfälle	201
1. Verkehrsunfallstatistik.....	201
2. Strafurteilsstatistik.....	203
3. Öffentliche Wahrnehmung der und Berichterstattung über „Raserunfälle“...	205
4. Das „Autorennen“ Urteil des Bundesgerichts (BGE 130 IV 58).....	207
§ 14 Rückschaufehler (<i>hindsight bias</i>)	209
I. Psychologie: Empirische Studien	211
II. Psychologie: Empirische Studien mit Experten	214
A. Ärzte	214
B. Buchprüfer	215
C. Richter	215
III. Recht: Empirische Studien	216
IV. Eigene Studie zum Rückschaufehler	219
A. Frage zum Rückschaufehler (2003)	219
B. Resultate (2003).....	219
C. Frage zum Rückschaufehler (2004)	220
D. Resultate (2004)	221
E. Diskussion	222
V. Rückschaufehler im Recht: Ignoranz oder Anpassung?	223
A. Rückschaufehler und privatrechtliche Verschuldenshaftung	224
1. Verschulden und Voraussehbarkeit.....	224
a. Voraussehbarkeit wird meist bejaht	225
b. Die Einhaltung von ex ante formulierten Vorschriften als Exkulpation ...	225
c. Die Problematik des Gefahrensatzes	227
2. Adäquater Kausalzusammenhang	228
§ 15 Hofeffekt (<i>halo effect</i>).....	230

Allgemeiner Teil

I.	Kritik an der Forschung zum Hofeffekt.....	232
II.	Hofeffekt und Strafzumessung.....	234
A.	Theoretisches Modell.....	234
B.	Experimente.....	236
C.	Feldstudien	238
III.	Eigene Studie zum Attraktivitäts-Hofeffekt	239
A.	Frage zum Attraktivitäts-Hofeffekt	239
B.	Resultate.....	240
C.	Diskussion	241
§ 16	Kontrast- und Kompromisseffekt	243
I.	Kontrast- und Anziehungseffekt	243
A.	Kontrasteffekt bei asymmetrisch dominierten Alternativen	244
B.	Anziehungseffekt bei klar minderwertigen Alternativen.....	246
II.	Kompromisseffekt (auch <i>extremeness aversion</i>)	248
III.	Beeinflussung rechtlicher Entscheidungen durch die Anzahl in Betracht gezogener Alternativen	250
A.	Klassische Studie von Neil Vidmar	250
B.	Demonstration von Kontrast- und Kompromisseffekt im Recht durch Kelman et al.....	252
IV.	Die lebenslängliche Verwahrung: eine neue Alternative	253
V.	Eigene Studie zum Kompromisseffekt	256
A.	Frage zum Kompromisseffekt	256
B.	Resultate.....	257
C.	Diskussion	258
1.	Erkenntnisse für den Gesetzgeber	259
2.	Erkenntnisse für den Praktiker	260
§ 17	Selbstüberschätzung und <i>overconfidence</i>	261
I.	Selbstüberschätzung	261
II.	Overconfidence	262
III.	Kulturelle und geschlechtliche Unterschiede bei Selbstüberschätzung und <i>overconfidence</i>	267
IV.	Overconfidence und Konfliktlösung	268

V. Eigene Studie zur Selbsteinschätzung.....	268
A. Frage zur Selbsteinschätzung	268
B. Resultate	268
C. Diskussion	269
§ 18 Zusammenfassung/Thesen	271
Anhang A: Herleitung des Bayes-Theorems aus den Axiomen der Wahrscheinlichkeits- theorie.....	275
Anhang B: Fragebogen 2003	278
Anhang C: Fragebogen 2004	282
Lebenslauf	286

Literaturverzeichnis

- ALEXANDER, GREGORY S.: Comparing the Two Legal Realisms – American and Scandinavian, *American Journal of Comparative Law* 2002, 131-174
- ALLAIS, MAURICE: Le comportement de l'homme rationnel devant le risque, critique des postulats et axiomes de l'école Américaine, *Econometrica* 1953, 503-546
- ANDERSON, CRAIG A./BUSHMAN, BRAD J.: Effects of Violent Video Games on Aggressive Behavior, Aggressive Cognition, Aggressive Affect, Physiological Arousal and Prosocial Behavior: A Meta-Analytic Review of the Scientific Literature, *Psychological Science* 2001, 353-359
- ANDERSON, JOHN C./JENNINGS, MARIANNE/LOWE, D. JORDAN/RECKERS, PHILIP M.J.: The Mitigation of Hindsight Bias in Judges' Evaluations of Auditor Decisions, *Auditing: Journal of Practice and Theory* 1997, 20-39
- /LOWE, D. JORDAN/RECKERS, PHILIP M.J.: Evaluation of auditor decisions: Hindsight bias effects and the expectation gap, *Journal of Economic Psychology* 1993, 711-737
- ANDERSON, JOHN R.: *Kognitive Psychologie*, 2. Auflage, Heidelberg 1996
- ARKES, HAL R./WORTMANN, ROBERT L./SAVILLE, PAUL D./HARKNESS, ALLAN R.: Hindsight Bias Among Physicians Weighing the Likelihood of Diagnoses, *Journal of Applied Psychology* 1981, 252-254
- ASCH, SOLOMON: Forming Impressions of Personality, *Journal of Abnormal and Social Psychology* 1946, 258-290
- AYTON, PETER/WRIGHT, GEORGE: Subjective Probability: What Should We Believe?, in: AYTON, PETER/WRIGHT, GEORGE (Hrsg.): *Subjective Probability*, New York 1994, 163-183
- BABCOCK, LINDA/FARBER, HENRY S./FOBIAN, CYNTHIA/SHAFIR, ELDAR: Forming Beliefs About Adjudicated Outcomes: Perceptions of Risk and Reservation Values, *International Review of Law and Economics* 1995, 289-303
- /LOEWENSTEIN, GEORGE/ISSACHAROFF, SAMUEL/CAMERER, COLIN: Biased Judgments of Fairness in Bargaining, *American Economic Review* 1995, 1337-1342
- /LOEWENSTEIN, GEORGE: Explaining Bargaining Impasse: The Role of Self-Serving Biases, *Journal of Economic Perspectives* 1997, 109-126
- BAKER, LYNN A./EMERY, ROBERT E.: When Every Relationship is Above Average: Perceptions and Expectations of Divorce at the Time of Marriage, *Law and Human Behavior* 1993, 439-450
- BALDING, DAVID J./DONNELLY, PETER: The Prosecutor's Fallacy and DNA Evidence, *Criminal Law Review* 1994, 711-721
- BALDWIN, MICHAEL R./KLEINKE, CHRIS L.: Effects of Severity of Accident, Intent, and "Alcoholism is a Disease" Excuse on Judgments of a Drunk Driver, *Journal of Applied Social Psychology* 1994, 2097-2109
- BÄR, WALTER: Zum Beweiswert der DNA-Analyse, in: DONATSCH, ANDREAS/SCHMID, NIKLAUS (Hrsg.): *Festschrift für Jörg Rehberg*, Zürich 1996, 41-51

- /KRATZER, ADELGUNDE: Die Anwendung der Gentechnologie in der Rechtsmedizin, in: Forschung für Leben Nr. 50, Mai 1998
- BARBER, BRAD M./ODEAN, TERRANCE: Boys Will Be Boys: Gender, Overconfidence, And Common Stock Investment, The Quarterly Journal of Economics 2001, 261-292
- BAR-HILLEL, MAYA: The Base-Rate Fallacy in Probability Judgments, Acta Psychologica 1979, 211-233
- Studies of Representativeness, in: KAHNEMAN/SLOVIC/TVERSKY (Hrsg.), 69-83
- /NETER, EFRAT: How Alike Is It? Versus How Likely Is It?: A Disjunction Fallacy in Probability Judgements, Journal of Personality and Social Psychology 1986, 1119 -1131; Nachdruck in: GILOVICH/GRIFFIN/KAHNEMAN (Hrsg.), 82-97
- BARON, JONATHAN/HERSHEY, JOHN C.: Outcome Bias in Decision Evaluations, Journal of Personality and Social Psychology 1989, 569-579
- /RITOV, ILANA: Reference points and omission bias, Organizational Behavior and Human Decision Making Processes 1994, 475-498
- BAUMEISTER, ROY F./DARLEY, JOHN F.: Reducing the Biasing Effect of Perpetrator Attractiveness in Jury Simulation, Personality and Social Psychology Bulletin 1982, 286-292
- BECKER, GARY S. : The Economics of Discrimination, Chicago 1957
- Crime and Punishment: An Economic Approach, Journal of Political Economy 1968, 169-217
- The Economic Approach to Human Behavior, Chicago 1976
- Treatise on the Family, Cambridge 1981
- /MURPHY, KEVIN M.: The Family and the State, Journal of Law and Economics 1988, 1-18
- BEDAU, HUGO/RADELET, MICHAEL: Miscarriages of Justice in Potentially Capital Cases, Stanford Law Review 1987, 21-90
- BENARTZI, SHLOMO/THALER, RICHARD H.: Myopic Loss Aversion and the Equity Premium Puzzle, Quarterly Journal of Economics 1995, 73-92
- BENASSI, MARIO A.: Effects of Order of Presentation, Primacy, and Physical Attractiveness on Attribution of Ability, Journal of Personality and Social Psychology 1982, 48-58
- BENDER, ROLF/NACK, ARMIN: Tatsachenfeststellung vor Gericht, Bd. 1 Glaubwürdigkeits- und Beweislehre, 2. Aufl. München 1995
- BENNETT, JONATHAN: Morality and Consequences, in: MCMURRIN, STERLING (Hrsg.): The Tanner Lectures on Human Values, Bd. 2, Salt Lake City 1981, 45-116
- BIERHOFF, HANS WERNER/BUCK, ERNST/KLEIN, RENATE: Attractiveness and Respectability of the Offender as Factors in the Evaluation of Criminal Cases, in: WEGENER, HERMANN/ LÖSEL, FRIEDRICH/HAISCH, JOCHEN (Hrsg.): Criminal Behavior and the Justice System: Psychological Perspectives, New York etc. 1989, 193-207
- BIRNBACHER, DIETER: Tun und Unterlassen, Stuttgart 1995

- BLANTON, HART/PELHAM, BRETT W./DEHART, TRACY/CARVALLO, MAURICIO: Overconfidence as Dissonance Reduction, *Journal of Experimental Social Psychology* 2001, 373-385
- BORNSTEIN, BRIAN H.: The Ecological Validity of Jury Simulations: Is the Jury Still Out?, *Law and Human Behavior* 1999, 75-92
- BOULOC BERNARD: *Droit pénal général*, 15. Aufl., Paris 1994
- BOURMISTROV-JÜTTNER, ELEONORA: *Subjektive Wahrscheinlichkeitstheorie und rationale Entscheidungstheorie in Anwendung auf die Rechtspraxis*, Diss. München 1987
- BRAITSTEIN, PAULA: Give P's a Chance: Understanding p-Values in Scientific Research, *Living+* Jan/Feb 2004; erhältlich unter www.bcpwa.org/articles/issue_28_34_give_ps_a_chance.pdf (besucht am 1. März 2005).
- BRAY, ROBERT M./KERR, NORBERT L.: Methodological Considerations in the Study of the Psychology of the Courtroom, in: KERR, NORBERT L./BRAY, ROBERT M. (Hrsg.): *The Psychology of the Courtroom*, San Diego etc. 1982, 287-323
- BREHM, ROLAND: *Kommentar zu Art. 41-61 OR*, 4. Aufl., Bern 1998
- BRENNER, LYLE A./KOEHLER, DEREK J./LIBERMAN, VARDA/TVERSKY, AMOS: Overconfidence in Probability and Frequency Judgments: A Critical Examination, *Organizational Behavior and Human Decision Processes* 1996, 212-219
- BRILMEYER, LEA/KORNHAUSER, LEWIS: Review: Quantitative Methods and Legal Decisions, *University of Chicago Law Review* 1978, 116-151
- BRODERSEN, KILIAN/ANSLINGER, KATJA/ROLF, BURKHARD: *DNA-Analyse und Strafverfahren*, München 2003
- BUDESCU, DAVID V./WALLSTEN, THOMAS S./WU, WING TUNG: On the Importance of Random Error in the Study of Probability Judgment. Part II: Applying the Stochastic Judgment Model to Detect Systematic Trends, *Journal of Behavioral Decision Making* 1997, 173-188
- BÜHLER, ALFRED: Beweiswürdigung, in: Leuenberger, Christoph: *Der Beweis im Zivilprozess*, Bern 2000, 71-92
- BURKE, DONALD M./AMES, M. ASHLEY/ETHERINGTON, ROSALIE/PIETSCH, JEAN: Effects of Victim's and Defendants Physical Attractiveness on the Perception of Responsibility in an Ambiguous Domestic Violence Case, *Journal of Family Violence* 1990, 199-207
- CALABRESI, GUIDO: Some Thoughts On Risk Distribution And The Law Of Torts, *Yale Law Journal* 1961, 499-533
- CAMERER, COLIN F.: Prospect Theory in the Wild: Evidence from the Field, in: KAHNEMAN/TVERSKY (Hrsg.), 288-300
- /HOGARTH, ROBYN M.: The Effects of Financial Incentives in Experiments: A Review and Capital-Labor Production Framework, *Journal of Risk and Uncertainty* 1999, 7-42
 - /ISSACHAROFF, SAMUEL/LOEWENSTEIN, GEORGE /O'DONOGHUE, TED/RABIN, MATTHEW: Regulation for Conservatives: Behavioral Economics and the Case for "Asymmetric Paternalism", *University of Pennsylvania Law Review* 2003, 1211-1254

- CARROLL, JOHN S./PAYNE, JOHN W.: Crime Seriousness, Recidivism Risk, and Causal Attributions in Judgments of Prison Term by Students and Experts, *Journal of Applied Social Psychology* 1977, 595-602
- CASPER, JONATHAN D./BENEDICT, KENNETTE/PERRY, JO L.: Juror Decision Making, Attitudes, and the Hindsight Bias, *Law and Human Behavior* 1989, 291-310
- CHAMPOD, CHRISTOPHE/TARONI, FRANCO: Probabilités au procès pénal – risques et solutions, *Schweizerische Zeitschrift für Strafrecht* 1994, 195-219
- CHAPMAN, GRETCHEN B./BORNSTEIN, BRIAN H.: The More You Ask For, the More You Get: Anchoring in Personal Injury Verdicts, *Applied Cognitive Psychology* 1996, 519-540
- CHAPMAN, LINDA/CHAPMAN, JOHN: Genesis of Popular but Erroneous Diagnostic Observations, *Journal of Abnormal Psychology* 1967, 193-204
- CHERTKOFF, JEROME M./CONLEY, MELINDA: Opening Offer and Frequency of Concessions as Bargaining Strategies, *Journal of Personality and Social Psychology* 1967, 181-185
- CHRISTENSEN-SZALANSKI, JAY J./BUSHYHEAD, JAMES B.: Physicians' Use of Probabilistic Information in Real Clinical Setting, *Journal of Experimental Psychology: Human Perception and Performance* 1981, 928-935
- /BEACH, LEE ROY: The Citation Bias: Fad and Fashion in the Judgment and Decision Literature, *American Psychologist* 1984, 75-78
 - /FOBIAN WILLHAM, CYNTHIA: The Hindsight Bias: A Meta-Analysis, *Organizational Behavior and Human Decision Processes* 1991, 147-168
- CHUNG, JANNE/MONROE, GARY S.: Gender Differences in Information Processing: An Empirical Test of the Hypothesis-Confirming Strategy in an Audit Context, *Accounting and Finance* 1998, 256-279
- CLOYD, C. BRYAN/SPILKER, BRIAN C.: The Influence of Client Preferences on Tax Professionals' Search for Judicial Precedents, Subsequent Judgments and Recommendations, *The Accounting Review* 1999, 299-322
- / - Confirmation Bias in Tax Information Search: A Comparison of Law Students and Accounting Students, *Journal of the American Taxation Association* 2000, 60-71
- COASE, RONALD: The Problem of Social Cost, *Journal of Law and Economics* 1960, 1-44
- Law and Economics at Chicago, *Journal of Law and Economics* 1993, 239-254
- COHEN, L. JONATHAN: *The Probable and the Provable*, London 1977
- Are There Any A Priori Constraints on the Study of Rationality, Behavioral and Brain Sciences 1981, 359-380
 - Subjective Probability and the Paradox of the Gatecrasher, *Arizona State Law Journal* 1981, 627-633
- COOPER, ARNOLD C./WOO, CAROLYN Y./DUNKELBERG, WILLIAM C.: Entrepreneurs' Perceived Chances for Success, *Journal of Business Venturing* 1988, 97-108
- COOPER, WILLIAM H.: Ubiquitous Halo, *Psychological Bulletin* 1981, 218-244
- COOTER, ROBERT/ULEN, THOMAS: *Law and Economics*, 4. Aufl., Reading 2003

- COSMIDES, LEDA/ TOOBY, JOHN: Are Humans Good Intuitive Statisticians After All? Rethinking Some Conclusions from the Literature on Judgment Under Uncertainty, *Cognition* 1996, 1-73
- COSTABILE, KRISTI A./KLEIN, STANLEY B.: Finishing Strong: Recency Effects in Juror Judgments, *Basic and Applied Social Psychology* 2005, 47-58
- CROSS, K. PATRICIA: Not Can, But Will College Teaching Be Improved, *New Directions for Higher Education* 1997, 1-15
- DANE, FRANCIS C./WRIGHTSMAN, LAWRENCE S.: Effects of Defendants' and Victims' Characteristics on Jurors' Verdicts, in: KERR, NORBERT L./BRAY, ROBERT M. (Hrsg.): *The Psychology of the Courtroom*, New York 1982, 83-115
- DARBY, BRUCE W./JEFFERS, DEVON: The Effects of Defendant and Juror Attractiveness on Simulated Courtroom Trial Decisions, *Social Behavior and Personality* 1988, 39-50
- DAVIS, DEBORAH/FOLLETTE, WILLIAM C.: Rethinking the Probative Value of Evidence: Base Rates, Intuitive Profiling, and the „Postdiction“ of Behavior, *Law and Human Behavior* 2002, 133-158
- DAWES, ROBYN M.: Representative Thinking in Clinical Judgment, *Clinical Psychology Review* 1986, 425-441
- /FAUST, DAVID/MEEHL, PAUL E.: Clinical versus Actuarial Judgment, in: GILOVICH/GRIFFIN/KAHNEMAN (Hrsg.), 716-729
 - /MULFORD, MATT: The False Consensus Effect and Overconfidence: Flaws in Judgment or Flaws in How We Study Them?, *Organizational Behavior and Human Decision Making Processes* 1996, 201-211
 - Probabilistic Thinking, in SMELSER, NEIL J./BALTES, PAUL B. (Hrsg.), *International Encyclopedia of the Social and Behavioral Sciences*, London 2001, 12082-12089
- DE FINETTI, BRUNO: *Theory of Probability, a Critical Introductory Treatment*, Chichester 1974
- DEFFENBACHER, KENNETH A.: Eyewitness Accuracy and Confidence: Can we Infer Anything about their Relationship?, *Law and Human Behavior* 1980, 243-260
- DHAMI, MANDEEP K./AYTON, PETER: Bailing and Jailing the Fast and Frugal Way, *Journal of Behavioral Decision Making* 2001, 141-168
- DION, KAREN/BERSCHIED, ELLEN/WALSTER, ELAINE: What Is Beautiful Is Good, *Journal of Personality and Social Psychology* 1972, 285-290.
- DIXIT, AVINASH K./NALEBUFF, BARRY J.: *Thinking Strategically*, New York/London 1993.
- DOHERTY, MICHAEL E./MYNATT, CLIFFORD R./TWENEY, RYAN D./SCHIAVO, MICHAEL D.: Pseudodiagnosticity, *Acta Psychologica* 1979, 111-121
- DONATSCH, ANDREAS: Die strafrechtlichen Grenzen der Sterbehilfe, in: METTNER, MATTHIAS (Hrsg.): *Wie menschenwürdig sterben?: Zur Debatte um die Sterbehilfe und die Praxis der Sterbebegleitung*, Zürich 2000, 121-137
- DONNELLY, PETER/FRIEDMAN, RICHARD D.: DNA Database Searches and the Legal Consumption of Scientific Evidence, *Michigan Law Review* 1997, 931-984

- DOYLE, JOHN R./O'CONNOR, DAVID J./REYNOLDS, GARETH M./BOTTOMLEY, PAUL A.: The Robustness of the Asymmetrically Dominated Effect: Buying Frames, Phantom Alternatives, and In-Store Purchases, *Psychology & Marketing* 1999, 225-243
- DROBNIG, ULRICH: Llewellyn and Germany, in: DROBNIG, ULRICH/ REHBINDER, MANFRED (Hrsg.): *Rechtsrealismus, multikulturelle Gesellschaft und Handelsrecht: Karl N. Llewellyn und seine Bedeutung heute*, Berlin 1994, 17-43
- EBBESEN, EBBE B./KONECNI, VLADIMIR J.: Decision Making and Information Integration in the Courts: The Setting of Bail, *Journal of Personality and Social Psychology* 1975, 805-821
- / - On the External Validity of Decision-Making Research: What Do We Know About Decisions in the Real World? in: WALLSTEN, THOMAS S. (Hrsg.): *Cognitive Processes in Choice and Decision Behavior*, Hillsdale 1980
- EDDY, DAVID M.: Probabilistic Reasoning in Clinical Medicine: Problems and Opportunities, in: KAHNEMAN/SLOVIC/TVERSKY (Hrsg.), 249-267
- EDWARDS, WARD/LINDMAN, HAROLD/SAVAGE, LEONHARD J.: Bayesian Statistical Inference for Psychological Research, *Psychological Review* 1963, 193-242
- Conservatism in human information processing, in: KLEINMUNTZ, B. (Hrsg.): *Formal Representation of Human Judgement*, New York 1968; gekürzter Nachdruck in: KAHNEMAN/SLOVIC/TVERSKY (Hrsg.), 359-369
- /VON WINTERFELDT, DETLOF: Cognitive Illusions and their Implications for the Law, *Southern California Law Review* 1986, 225-276
- EFRAN, MICHAEL G.: The Effect of Physical Appearance on the Judgment of Guilt, Interpersonal Attraction, and Severity of Recommended Punishment in a Simulated Jury Task, *Journal of Research in Personality* 1974, 45-54
- FREEMAN, M.D.A.: *Lloyd's Introduction to Jurisprudence*, 7. Aufl. London 2001
- EGLI, URS: *Vergleichsdruck im Zivilprozess*, Berlin 1996
- EHRlich, EUGEN: *Freie Rechtsfindung und freie Rechtswissenschaft*, Leipzig 1903, Nachdruck in: EHRlich, EUGEN: *Recht und Leben: Gesammelte Schriften zur Rechtstatistik, sachenforschung und zur Freirechtslehre, ausgewählt und eingeleitet von MANFRED REHBINDER*, Berlin 1967, 170-202
- EIDENMÜLLER, HORST: *Effizienz als Rechtsprinzip*, Tübingen 1995
- EISENBERG, MELVIN A.: The Limits of Cognition and the Limits of Contract, *Stanford Law Review* 1995, 211-259
- EISENFÜHR, FRANZ/WEBER, MARTIN: *Rationales Entscheiden*, Heidelberg 1993
- EKELÖF, PER OLOF: Beweiswert, in: GRUNSKY, WOLFGANG/STÜMER, ROLF/WALTER, GERHARD/ WOLF, MANFRED (Hrsg.): *Festschrift für Fritz Baur*, Tübingen 1981, 343-363
- ELFFERS, HENK/HESSING, DICK J.: Influencing the Prospects of Tax Evasion, *Journal of Economic Psychology* 1997, 289-304
- ELlickson, ROBERT C.: Bringing Culture and Human Frailty to Rational Actors: A Critique of Classical Law and Economics, *Chicago-Kent Law Review* 1989, 23- 55
- Order Without Law, Cambridge 1991

- ELWORK, AMIRAM/SALES, BRUCE DENNIS/SUGGS, DAVID: The Trial: A Research Review, in: SALES (Hrsg.), 1-59
- EMBLY, CRAIG/GELARDI, ALEXANDER M. G./LOWE, D. JORDAN: A Research Note on the Influence of Outcome Knowledge on Audit Partners' Judgments, Behavioral Research in Accounting 2002, 86-103
- ENGLISH, BIRTE/MUSSWEILER, THOMAS: Sentencing Under Uncertainty: Anchoring Effects in the Court Room, Journal of Applied Social Psychology 2001, 1535-1551
- EREV, IDO/WALLSTEN, THOMAS S./BUDESCU, DAVID V.: Simultaneous Over- and Underconfidence: The Role of Error in Judgment Processes, Psychological Review 1994, 519-527
- EVANS, JONATHAN ST. B. T.: Bias in Human Reasoning: Causes and Consequences, Hove 1989
- /VENN, SIMON/FEENEY, AIDAN: Implicit and Explicit Processes in a Hypothesis Testing Task, British Journal of Psychology 2002, 31-46
- FAHRMEIER, LUDWIG/KÜNSTLER, RITA/PIGEOT, IRIS/KUNZ, GERHARD: Statistik, 5. Aufl. Berlin etc. 2004
- FAIGMAN, DAVID L./BAGLIONI, A.J.: Bayes' Theorem in the Trial Process: Instructing Jurors on the Value of Statistical Evidence, Law and Human Behavior 1988, 1-17
- FARBER, DANIEL A.: Towards A New Legal Realism, University of Chicago Law Review 2001, 279-303
- FARIÑA, FRANCISCA/ARCE, RAMON/NOVO, MERCEDES: Anchoring in Judicial Decision-Making, Psychology in Spain 2003
- FEENEY, AIDAN/EVANS, JONATHAN ST.B.T./CLIBBENS, JOHN: Background beliefs and evidence interpretation, Thinking and Reasoning 2000, 97-124
- /EVANS, JONATHAN ST.B.T./VENN, SIMON: A Rarity Heuristic for Hypothesis Testing, in: GLEITMAN, LILA R./JOSHI, ARAVID K. (Hrsg.), Proceedings of the 22nd Annual Conference of the Cognitive Science Society, Mahwah 2000, 119-124
- FEHR, ERNST/ SCHWARZ, GERHARD (Hrsg.): Psychologische Grundlagen der Ökonomie, Zürich 2002
- /FISCHBACHER, URS/GÄCHTER, SIMON: Strong Reciprocity, Human Cooperation and the Enforcement of Social Norms, Human Nature 2002, 1-25
- /FISCHBACHER, URS: The Nature of Human Altruism, Nature 2003, 785-791
- FEINBERG, JOEL: The Moral Limits of Criminal Law, Vol. 1: Harm to Others, New York 1984
- FEINGOLD, ALAN: Good-looking People Are Not What We Think, Psychological Bulletin 1992, 304-341
- FEZER, KARL-HEINZ: Aspekte einer Rechtskritik an der economic analysis of law und am property rights approach, Juristen-Zeitung 1986, 817-864
- FIEDLER, KLAUS: Beruhen Bestätigungsfehler nur auf einem Bestätigungsfehler?, Psychologische Beiträge 1983, 280-286
- FIELDING, JANE/GILBERT, NIGEL: Understanding Social Statistics, London 2000

- FINKELSTEIN, MICHAEL/FAIRLEY, WILLIAM: A Bayesian Approach to Identification Evidence, *Harvard Law Review* 1970, 489-517
- /- A Comment on „Trial by Mathematics“, *Harvard Law Review* 1971, 1801-1809
- FISCHHOFF, BARUCH: Hindsight \neq Foresight: The Effect of Outcome Knowledge on Judgment Under Uncertainty, *Journal of Experimental Psychology* 1975, 288-299
- /SLOVIC, PAUL/LICHTENSTEIN, SARAH: Knowing with Certainty: The Appropriateness of Extreme Confidence, *Journal of Experimental Psychology: Human Perception and Performance* 1977, 552-564
- Perceived Informativeness of Facts, *Journal of Experimental Psychology: Human Perception and Performance* 1978, 349-358
- For Those Condemned to Study the Past: Heuristics and Biases in Hindsight, in: KAHNEMAN/SLOVIC/TVERSKY (Hrsg.), 335-354
- FISHER, ROGER/URY, WILLIAM/PATTON, BRUCE: *Getting to Yes*, 2. Aufl., New York 1991
- FITZMAURICE, CATHERINE/PEASE, KEN: *The Psychology of Judicial Sentencing*, Manchester 1986
- FRANK, JEROME: *Law and the Modern Mind*, New York 1930
- *Courts on Trial: Myth and Reality in American Justice*, Princeton 1949
- FRANK, RICHARD/STRÄULI, HANS/MESSMER, GEORG: *Kommentar zur zürcherischen Zivilprozessordnung*, 3. Aufl. Zürich 1997
- FRIEDMAN, MILTON: The Methodology of Positive Economics, in: FRIEDMAN, MILTON: *Essays in Positive Economics*, Chicago 1953, 3-43
- FRIEDMANN, RICHARD D./PARK, ROGER C.: Sometimes What Everybody Thinks They Know is True, *Law and Human Behavior* 2003, 629-644
- FRIEDRICH, JAMES: Primary Error Detection and Minimization (PEDMIN) Strategies in Social Cognition: A Reinterpretation of Confirmation Bias Phenomena, *Psychological Review* 1993, 298-319
- The “Bias” Bias in Social Psychology: Adaptive When and How?, *Behavioral and Brain Sciences* 2004, 335-336
- FRIEND, RONALD M./VINSON, MICHAEL: Leaning Over Backwards: Jurors’ Responses to Defendants’ Attractiveness, *Journal of Communication* 1974, 124-129
- FUNDER, DAVID C.: On the Accuracy of Personality Judgment: a Realistic Approach, *Psychological Review* 1995, 652-670
- GADENNE, VOLKER: Der Bestätigungsfehler und die Rationalität kognitiver Prozesse, *Psychologische Beiträge* 1982, 11-25
- /OSWALD, MARGRIT: Entstehung und Veränderung von Bestätigungstendenzen beim Testen von Hypothesen, *Zeitschrift für experimentelle und angewandte Psychologie* 1986, 360-374
- Deduktives Denken und Rationalität, in: HELL/FIEDLER/GIGERENZER (Hrsg.), 161-188
- GALANTER, MARC/EDWARDS, MARK ALAN, Introduction: The Path of the Law Ands, *Wisconsin Law Review* 1997, 375-387

Allgemeiner Teil

- GALINSKY, ADAM D./MUSSWEILER, THOMAS: First Offers as Anchors: The Role of Perspective-Taking and Negotiator Focus, *Journal of Personality and Social Psychology* 2001, 657-669
- /MEDVEC, VICTORIA H.: Disconnecting Outcomes and Evaluations: The Role of Negotiator Focus, *Journal of Personality and Social Psychology* 2002, 1131-1140
- /SEIDEN, VANESSA L./KIM, PETER H./MEDVEC, VICTORIA H.: The Dissatisfaction of Having Your First Offer Accepted: The Role of Counterfactual Thinking in Negotiations, *Personality and Social Psychology Bulletin* 2002, 271-283
- GALTON, FRANCIS: Regression Towards Mediocrity in Hereditary Stature, *Journal of the Anthropological Institute* 1886, 246-63
- GARRIDO MARTIN, EUGENIO/HERRERO ALONSO, CARMEN: Influence of the Prosecutor's Plea on the Judge's Sentencing in Sexual Crimes: Hypothesis of the Theory of Anchoring by Tversky and Kahneman, in: REDONDO, SANTIAGO/GARRIDO, VICENTE/PÉREZ, JORGE / BARBERET, ROSEMARY: *Advances in Psychology and Law*, Berlin etc. 1997, 215-226
- GENTILE, DOUGLAS A./WALSH, DAVID A.: A Normative Study of Family Media Habits, *Applied Developmental Psychology* 2002, 157-178
- GIGERENZER, GERD: How to Make Cognitive Illusions Disappear: Beyond "Heuristics and Biases", *European Review of Social Psychology* 1991, 83-115
- /HOFFRAGE, ULRICH/KLEINBÖLTING, HANS: Probabilistic Mental Models: A Brunswikian Theory of Confidence, *Psychological Review* 1991, 506-528
- Why the Distinction between Single-event Probabilities and Frequencies is Important for Psychology (and Vice Versa), in: WRIGHT/AYTON (Hrsg.), 129-161
- Die Repräsentation von Information und ihre Auswirkung auf statistisches Denken, in: HELL/FIEDLER/GIGERENZER (Hrsg.), 99-127
- /HOFFRAGE, ULRICH: How to Improve Bayesian Reasoning Without Instructions: Frequency Formats, *Psychological Review* 1995, 684-704
- On Narrow Norms and Vague Heuristics: A Rebuttal to Kahneman and Tversky, *Psychological Review* 1996, 592-596
- Ecological Intelligence: An Adaptation for Frequencies, in: CUMMINS, DENISE D./ALLEN, COLIN (Hrsg.): *The Evolution of Mind*, New York 1998, 9-29
- /HOFFRAGE, ULRICH: Overcoming Difficulties in Bayesian Reasoning: A Reply to Lewis and Keren (1999) and Mellers and McGraw (1999), *Psychological Review* 1999, 425-430
- /CZERLINSKI, JEAN/MARTIGNON, LAURA: How Good Are Fast and Frugal Heuristics?, in: GILOVICH/GRIFFIN/KAHNEMAN (Hrsg.), 559-581
- GILOVICH, THOMAS/MEDVEC, VICTORIA H.: The Experience of Regret: What, When, and Why, *Psychological Review* 1995, 379-395
- /GRIFFIN, DALE/KAHNEMAN, DANIEL (Hrsg.): *Heuristics and Biases: The Psychology of Intuitive Judgment*, Cambridge 2002 (zit.: GILOVICH/GRIFFIN/ KAHNEMAN (Hrsg.))
- GOETTE, LORENZ/FEHR, ERNST/HUFFMANN, DAVID: Loss Aversion and Labor Supply, *Journal of the European Economic Association* 2004, 216-228

- GOLDBERG, LEWIS R.: The Effectiveness of Clinicians' Judgments: The Diagnosis of Organic Brain Damage from the Bender-Gestalt Test, *Journal of Consulting Psychology* 1959, 25-33
- GOOD, IRVING J.: *Probability and the Weighing of Evidence*, London 1950
- When Batterer Turns Murderer, *Nature* 1995, 541
- GOODIE, ADAM S./FANTINO, EDMUND: An Experientially Derived Base-Rate Error in Humans, *Psychological Science* 1995, 101-106
- GOODMAN-DELAHUNTY, JANE/GRANHAG, PAR-ANDERS/LOFTUS, ELIZABETH F.: How Well Can Lawyers Predicts Their Chances of Success?, U. of Washington 1998 (nicht publiziert)
- GRÄNS, MINNA: *Das Risiko materiell fehlerhafter Urteile*, Diss. Uppsala 1995, München 2002
- GREGER, REINHARD: *Beweis und Wahrscheinlichkeit*, Diss. Erlangen-Nürnberg, München 1978
- GRIFFIN, DALE/TVERSKY, AMOS: The Weighing of Evidence and the Determinants of Confidence, *Cognitive Psychology* 1992, 411-435; Nachdruck in: GILOVICH/GRIFFIN/KAHNEMAN (Hrsg.), 230-249. Zitiert wird der Nachdruck.
- GROFMAN, BERNARD: Mathematical Models of Juror and Jury Decision-Making, in: SALES, BRUCE DENNIS (Hrsg.): *The Trial Process*, New York 1981, 305-352
- GROSS, SAMUEL R./SYVERUD, KENT D.: Getting to No: A Study of Settlement Negotiations and the Selection of Cases for Trial, *Michigan Law Review* 1991, 319-393
- GUIGNARD, JEAN-PIERRE: Note sur l'arrêt X., *JdT* 1988, 131
- GUILBAULT, REBECCA L./BRYANT, FRED B./HOWARD BROCKWAY, JENNIFER/POSAVAC, EMIL J.: A Meta-Analysis of Research on Hindsight Bias, *Basic and Applied Social Psychology* 2004, 103-117
- GULDENER, MAX: *Beweiswürdigung und Beweislast nach schweizerischem Zivilprozessrecht*, Zürich 1955
- *Schweizerisches Zivilprozessrecht*, 3. Aufl., Zürich 1979
- GUTHRIE, CHRIS: Framing Frivolous Litigation: A Psychological Theory, *University of Chicago Law Review* 2000, 163-216
- /RACHLINSKI, JEFFREY J./WISTRICH, ANDREW J.: Inside the Judicial Mind, *Cornell Law Review* 2001, 777-830.
 - Prospect Theory, Risk Preference, And The Law, *Northwestern University Law Review* 2003, 1115-1163
- HABSCHEID, WALTER: *Schweizerisches Zivilprozess- und Gerichtsorganisationsrecht*, 2. Aufl. Zürich 1990
- HAIDT, JONATHAN/BARON, JONATHAN: Social Roles and the Moral Judgement of Acts and Omission, *European Journal of Social Psychology* 1996, 201-218
- HAISCH, JOCHEN: Verhaltensursachen bei Straftätern und Strafurteile durch Juristen und Laien, *Monatsschrift für Kriminologie* 1977, 82-88

- HANSON, JON D./KYSAR, DOUGLAS A.: Taking Behavioralism Seriously: Some Evidence of Market Manipulation, *Harvard Law Review* 1999, 1420-1572
- /- Taking Behavioralism Seriously: The Problem of Market Manipulation, *New York University Law Review* 1999, 630-749
- HARLEY, ERIN M./CARLSEN, KERI A./LOFTUS, GEOFFREY R.: The "Saw-It-All-Along" Effect: Demonstrations of Visual Hindsight Bias, *Journal of Experimental Psychology: Learning, Memory and Cognition* 2004, 960-968
- HARVEY, NIGEL: Relations Between Confidence and Skilled Performance, in: WRIGHT/AYTON (Hrsg.), 321-352
- HASTIE, REID (Hrsg.): *Inside the Juror*, Cambridge 1993
- /VISCUSI, W. KIP: What Juries Can't Do Well: The Jury's Performance as a Risk Manager, *Arizona Law Review* 1998, 901-921
- /SCHADKE, DAVID A./PAYNE, JOHN W.: Juror Judgements in Civil Cases: Effects of Plaintiff's Requests and Plaintiff's Identity on Punitive Damage Awards, *Law and Human Behavior* 1999, 445-469
- /-/- Juror Judgments in Civil Cases: Hindsight Effects on Judgments of Liability for Punitive Damages, *Law and Human Behavior* 1999, 597-613
- /DAWES, ROBYN M.: *Rational Choice in an Uncertain World*, Thousand Oaks 2001
- Algebraic Models of Juror Decision Processes, in: HASTIE (Hrsg.), 84-115
- HATAKKA, MINNA/TAPANI KLAMI, HANNU: *Beweismass und Irrtumsrisiko*, Helsinki 1990
- HAWKINS, SCOTT A./HASTIE, REID: Hindsight: Biased Judgments of Past Events After the Outcomes Are Known, *Psychological Bulletin* 1990, 311-327
- HEATH, CHIP/LARRICK, RICHARD P./WU, GEORGE: Goals as Reference Points, *Cognitive Psychology* 1999, 79-109
- HELL, WOLFGANG/FIEDLER, KLAUS /GIGERENZER, GERD (Hrsg.), *Kognitive Täuschungen: Fehl-Leistungen und Mechanismen des Urteilens, Denkens und Erinnerns*, Heidelberg 1993
- Kognitive und optische Täuschungen, in: HELL, WOLFGANG/FIEDLER, KLAUS /GIGERENZER, GERD (Hrsg.), 317-324
- HENRICH, JOSEPH/BOYD, ROBERT/BOWLES, SAMUEL /CAMERER, COLIN/FEHR, ERNST/GINTIS, HERBERT: *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*, New York 2004
- HERGET, JAMES E./WALLACE, STEPHEN: The German Free Law Movement as the Source of American Legal Realism, *Vanderbilt Law Review* 1987, 399-455.
- HERTIG, GÉRARD: Switzerland, *International Review of Law and Economics* 1991, 293-300
- HERTWIG, RALPH/GIGERENZER, GERD/HOFFRAGE, ULRICH: The Reiteration Effect in Hindsight Bias, *Psychological Review* 1997, 194-202
- HETTICH, WALTER/WINER, STANLEY L.: A positive model of tax structure, *Journal of Public Economics* 1984, 67-87
- HINSZ, VERLIN B./INDAHL, KRISTIN E.: Assimilation to Anchors for Damage Awards in a Mock Civil Trial, *Journal of Applied Social Psychology* 1995, 991-1026

- HITCHCOCK, CHRISTOPHER: Probability and Chance: Philosophical Aspects, in: SMELSER/BALTES (Hrsg.), 12089-12095
- HOFFMAN, ELIZABETH/SPITZER, MATTHEW L.: Willingness to Pay vs. Willingness to Accept: Legal and Economic Implications, Washington University Law Quarterly 1993, 59-114
- HOFFRAGE, ULRICH: Die Illusion der Sicherheit bei Entscheidungen unter Unsicherheit, in: HELL/FIEDLER/GIGERENZER (Hrsg.), 73-97
- /GIGERENZER, GERD: Using Natural Frequencies to Improve Diagnostic Inferences, Academic Medicine 1998, 538-54
 - /HERTWIG, RALPH/GIGERENZER, GERD: Hindsight Bias: A By-Product of Knowledge Updating?, Journal of Experimental Psychology: Learning, Memory and Cognition 2000, 566-581
 - /POHL, RÜDIGER F.: Research on Hindsight Bias: A Rich Past, a Productive Present and a Challenging Future, Memory 2003, 329-335
- HOGARTH, ROBIN M./EINHORN, HILLEL J.: Order Effects in Belief Updating: The Belief-Adjustment Model, Cognitive Psychology 1992, 1-55
- HOLMES, JR., OLIVER WENDELL: The Path of the Law, Harvard Law Review 1897, 457-469
- HONSELL, HEINRICH: Schweizerisches Haftpflichtrecht, 3. Aufl., Zürich 2000
- HOROWITZ, JOHN K./MCCONNELL, KENNETH E.: A Review of WTA/WTP Studies, Journal of Environmental Economics and Management 2002, 426-447
- HOTHER, MARC: Die DNA-Analyse – ihre Bedeutung für die Strafverfolgung und ihr Beweiswert im Strafverfahren, Diss. Giessen 1995
- HUBER, JOEL/PAYNE, JOHN W./PUTO, CHRISTOPHER: Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis, Journal of Consumer Research 1982, 90-98
- /PUTO, CHRISTOPHER: Market Boundaries and Product Choice: Illustrating Attraction and Substitution Effects, Journal of Consumer Research 1983, 31-44
- HÜTTE, KLAUS/DUCKSCH, PETRA: Die Genugtuung, Zürich 2003
- IZZETT, RICHARD R./LEGINSKI, WALTER: Group Discussion and the Influence of Defendant Characteristics in a Simulated Jury Setting 1974, 271-279
- JACOBSON, STEVEN K./BERGER, CHARLES R.: Communication And Justice: Defendant Attributes And Their Effects on the Severity of His Sentence, Speech Monographs 1974, 282-286
- JANSON, GUNNAR: Ökonomische Theorie im Recht, Berlin 2004
- JEFFREY, RICHARD: Subjective Probability: The Real Thing, Cambridge 2004
- JOHNSON, ERIC J./HERSHEY, JOHN/MEZAROS, JACQUELINE/KUNREUTHER, HOWARD: Framing, Probability Distortions, and Insurance Decisions, Journal of Risk and Uncertainty 1993, 35-51
- JOHNSON, LINDA M.: An Empirical Investigation of the Effects of Advocacy on Preparers' Evaluations of Judicial Evidence, Journal of the American Taxation Association 1993, 1-22

Allgemeiner Teil

- JOLLS, CHRISTINE/SUNSTEIN, CASS R./THALER, RICHARD H.: A Behavioral Approach to Law and Economics, *Stanford Law Review* 1998, 1471-1550
- JONAS, EVA/SCHULZ-HARDT, STEFAN/FREY, DIETER: Konfirmatorische Informationssuche bei simultaner vs. sequentieller Informationsvorgabe, *Zeitschrift für Experimentelle Psychologie* 2001, 239-247
- JONES, OWEN D.: Time-Shifted Rationality and the Law of Law's Leverage: Behavioral Economics Meets Behavioral Biology, *Northwestern University Law Review* 2001, 1141-1205
- JUNGERMANN, HELMUT/PFISTER, HANS-RÜDIGER/FISCHER, KATRIN: *Die Psychologie der Entscheidung*, Heidelberg 1998
- JUSLIN, PETER: The Overconfidence Phenomenon as a Consequence of Informal Experimenter Guided Selection of Almanac Items, *Organizational Behavior and Human Decision Processes* 1994, 226-246
- KACHELMEIER, STEVEN J./SHEHATA, MOHAMED: Examining Risk Preferences Under High Monetary Incentives: Experimental Evidence From the People's Republic of China, *American Economic Review* 1992, 1120-1141
- KADOUS, KATHRYN: The Effects of Audit Quality and Consequence Severity on Juror Evaluations of Auditor Responsibility for Plaintiff Losses, *The Accounting Review* 2000, 327-341
- KAGEHIRO, DOROTHY K./TAYLOR, RALPH B./LAUFER, WILLIAM S./HARLANDT, ALAN T.: Hindsight Bias and Third-Party Consentors to Warrantless Police Searches, *Law and Human Behavior* 1991, 305-314
- KAHNEMAN, DANIEL/TVERSKY, AMOS: Subjective probability: A judgment of representativeness, *Cognitive Psychology* 1972, 430-454
- /- On the Psychology of Prediction, *Psychological Review* 1973, 237-251; Nachdruck in: KAHNEMAN/SLOVIC/TVERSKY (Hrsg.), 48-68
- /- Prospect theory: An Analysis of Decision Under Risk, *Econometrica* 1979, 263-291; Nachdruck in: KAHNEMAN/TVERSKY (Hrsg.), 17-43. Zitiert wird der Nachdruck.
- /SLOVIC, PAUL/TVERSKY, AMOS (Hrsg.), *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge 1982 (zitiert als: KAHNEMAN/ SLOVIC/TVERSKY (Hrsg.))
- /- The Simulation Heuristic, in: KAHNEMAN/SLOVIC/TVERSKY (Hrsg.), 201-208
- /- On the Study of Statistical Intuitions, *Cognition* 1982, 123-141
- /- Choices, Values, and Frames, *American Psychologist* 1984, 341-350, Nachdruck in: KAHNEMAN/TVERSKY (Hrsg.), 1-16
- /KNETSCH, JACK L./THALER, RICHARD H.: Experimental Tests of the Endowment Effect and the Coarse Theorem, *Journal of Political Economy* 1990, 1325-1348, Nachdruck in: SUNSTEIN (Hrsg.), 211-231
- /-/- The Endowment Effect, Loss Aversion and the Status Quo Bias, *Journal of Economic Perspectives* 1991, 193-206
- New Challenges to the Rationality Assumption, *Journal of institutional and theoretical economics* 1994, 18-36

- Varieties of Counterfactual Thinking, in: ROESE, NEAL J./OLSON, JAMES M.: What Might Have Been: The Social Role of Counterfactual Thinking, New York 1995, 375-396
- /TVERSKY, AMOS: On the Reality of Cognitive Illusions, Psychological Review 1996, 582-591
- /- Choices, Values and Frames, Cambridge 2000 (zitiert: KAHNEMAN/TVERSKY (Hrsg.))
- /FREDERICK, SHANE: Representativeness Revisited: Attribute Substitution in Intuitive Judgment, in: GILOVICH/GRIFFIN/KAHNEMAN (Hrsg.), 49-81
- KAMIN, KIM A./RACHLINSKI, JEFFREY: Ex Post \neq Ex Ante: Determining Liability in Hindsight, Law and Human Behavior 1995, 89-104
- KANAZAW, SATOSHI/KOVAR, JODY L.: Why Beautiful People Are More Intelligent, Intelligence 2004, 227-243
- KASSIN, SAUL M./GOLDSTEIN, CHRISTINE C./SAVITSKY, KENNETH: Behavioral Confirmation in the Interrogation Room: On the Dangers of Presuming Guilt, Law and Human Behavior 2003, 187-203
- KATYAL, NEAL KUMA: Deterrence's Difficulty, Michigan Law Review 1997, 2385-2476
- KAYE, DAVID H.: The Laws of Probability and the Laws of the Land, University of Chicago Law Review 1979, 34-56
- DNA Evidence: Probability, Population Genetics, and the Courts, Harvard Journal of Law & Technology 1993, 101-172
- /KOEHLER, JONATHAN J.: The Misquantification of Probative Value, Law and Human Behavior 2003, 645-659
- KELLER, ALFRED: Haftpflicht im Privatrecht, 6. Aufl., Bern 2002
- KELMAN, MARK/ROTTENSTREICH, YUVAL/TVERSKY, AMOS: Context-Dependence in Legal Decision Making, Journal of Legal Studies 1996, 287-318, Nachdruck in: SUNSTEIN (Hrsg.), 61-94. Zitiert wird der Nachdruck.
- Behavioral Economics as Part of a Rhetorical Duet: A Response to Jolls, Sunstein and Thaler, Stanford Law Review 1998, 1577-1591
- Law and Behavioral Science: Conceptual Overview, Northwestern University Law Review 2003, 1347-1392
- KERR, NORBERT L./BRAY, ROBERT M. (Hrsg.): The Psychology of the Courtroom, San Diego etc. 1982
- KIRCHGÄSSNER, GEBHARD: Homo Oeconomicus, 2. Aufl., Tübingen 2000
- KIRCHLER, ERICH/MACIEJOVSKY, BORIS: Steuermoral und Steuerhinterziehung, in: FREY, DIETER / VON ROSENSTIEL, LUTZ (Hrsg.): Enzyklopädie der Psychologie – Wirtschaftspsychologie, Göttingen 2002
- KIRCHNER, CHRISTIAN: The Difficult Reception of Law and Economics in Germany, International Review of Law and Economics 1991, 277-292
- KIVETZ, RAN/NETZER, ODED/SRINIVASAN, V.: Alternative Models for Capturing the Compromise Effect, Journal of Marketing Research 2004, 237-257

Allgemeiner Teil

- KLAUER, KARL CHRISTOPH/SCHMELING, ARMIN: Sind Halo-Fehler Flüchtigkeitsfehler?, Zeitschrift für experimentelle und angewandte Psychologie 1990, 594-607
- KLAYMAN, JOSHUA/HA, YOUNG-WON: Confirmation, Disconfirmation, and Information in Hypothesis Testing, Psychological Review 1987, 211-228
- Varieties of Confirmation Bias, in: BUSEMEYER, JEROME/HASTIE, REID/ MEDIN, DOUGLAS L. (Hrsg.): Decision Making from a Cognitive Perspective, New York 1995
 - /BELL, JACK S./GONZALEZ-VALLEJO, CLAUDIA/BARLAS, SEMA: Overconfidence: It Depends on How, What, and Whom You Ask, Organization Behavior and Human Decision Processes 1999, 216-247
- KLEITER, GERNOT D./KREBS, MARIANNE: Do Subjects Understand Base Rates?, Organizational Behavior and Human Decision Processes 1997, 25-61
- KNIGHT, FRANK H.: Risk, Uncertainty and Profit, Boston 1921
- KOCH, CHANTAL MEES/DEVINE, DENNIS J.: Effects of Reasonable Doubt Definition and Inclusion of a Lesser Charge on Jury Verdicts, Law and Human Behavior 1999, 653-674
- KOEHLER, DEREK J.: Explanation, Imagination, and Confidence in Judgment, Psychological Bulletin 1991, 499-519
- /BRENNER, LYLE/GRIFFIN, DALE: The Calibration of Expert Judgment: Heuristics and Biases Beyond the Laboratory, in: GILOVICH/GRIFFIN/KAHNEMAN (Hrsg.), 686-715
- KOEHLER, JONATHAN J./SHAVIRO, DANIEL N.: Veridical Verdicts: Increasing Verdict Accuracy Through the Use of Overtly Probabilistic Evidence and Methods, Cornell Law Review 1990, 247-279
- The Influence of Prior Beliefs on Scientific Judgements of Evidence Quality, Organizational Behavior and Human Decision Making Processes 1993, 28-55
 - /CHIA, AUDREY/LINDSEY, J. SAM: The Random Match Probability (RMP) in DNA Evidence: Irrelevant and Prejudicial?, Jurimetrics 1995, 201-219
 - The Base Rate Fallacy Reconsidered: Descriptive, Normative and Methodological Challenges, Behavioral and Brain Sciences 1996, 1-53
 - One in Millions, Billions and Trillions: Lessons from People v. Collins (1968) for People v. Simpson (1995), Journal of Legal Education 1997, 214-223
 - Why DNA Likelihood-Ratios Should Account for Error (Even When A National Research Council Report Says They Should Not), Jurimetrics 1997, 425-437
- KONECNI, VLADIMIR J./EBBESSEN, EBBE B.: A Critique of Theory and Method in Social-Psychological Approaches to Legal Issues, in: SALES, BRUCE DENNIS (Hrsg.), The Trial Process, New York 1981, 481-498
- /- (Hrsg.): The Criminal Justice System: A Social-Psychological Analysis, San Francisco 1982
 - /- The Setting of Bail, in: KONECNI/EBBESSEN (Hrsg.), 191-230
 - /- The Mythology of Legal Decision Making, International Journal of Law and Psychiatry 1984, 5-18

- /- Methodische Probleme in der Forschung über juristische Entscheidungsprozesse – unter besonderer Berücksichtigung experimenteller Simulationen, Gruppendynamik 1991, 175-188.
- KORDES-DE VAAL, JOHANNA H.: Intention and the omission bias: Omissions perceived as nondecisions, *Acta Psychologica* 1996, 161-172
- KORIAT, ASHER/LICHTENSTEIN, SARAH/FISCHHOFF, BARUCH: Reasons for Confidence, *Journal of Experimental Psychology: Human Learning and Memory* 1980, 107-118
- KOROBKIN, RUSSELL: Policymaking and the Offer/Asking Price Gap: Toward a Theory of Efficient Entitlement Allocation, *Stanford Law Review* 1994, 663-708
- /GUTHRIE, CHRIS: Opening Offers and Out-of-Court Settlement: A Little Moderation May not Go a Long Way, *Ohio State Journal on Dispute Resolution* 1994, 1-20
- /- Psychological Barriers to Litigation Settlement: An Experimental Approach, *Michigan Law Review* 1994, 107-192
- Behavioral Economics, Contract Formation, and Contract Law, in: SUNSTEIN (Hrsg.), 116-143
- Inertia and Preference in Contract Negotiation: The Psychological Power of Default Rules and Form Terms, *Vanderbilt Law Review* 1998, 1583-1651
- The Status Quo Bias and Contract Default Rules, *Cornell Law Review* 1998, 608-687
- /ULEN, THOMAS S.: Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics, *California Law Review* 2000, 1053-1144
- Aspirations and Settlement, *Cornell Law Review* 2002, 1-61
- KRÄMER, WALTER: Statistik verstehen, 4. Aufl. München 2004
- KRAUS, STEFAN/HERTWIG, RALPH: Muss DNA-Evidenz schwer verständlich sein?, *Monatsschrift für Kriminologie und Strafrechtsreform* 2000, 155-162
- KRAY, LAURA J./GALINSKY, ADAM D.: The Debiasing Effect of Counterfactual Mind-Sets: Increasing the Search for Disconfirmatory Information in Group Decisions, *Organizational Behavior and Human Decision Processes* 2003, 69-81
- KRISTENSEN, HENRIK/GÄRLING, TOMMY: Anchoring Induced Biases in Consumer Price Negotiations, *Journal of Consumer Policy* 2000, 445-460
- KRUEGER, JOACHIM L./FUNDER, DAVID C.: Towards a Balanced Social Psychology: Causes, Consequences, and Cures for the Problem Seeking Approach to Social Behavior and Cognition, *Behavioral and Brain Sciences* 2004, 313-376
- KÜHBERGER, ANTON/SCHULTE-MECKLENBECK, MICHAEL/PERNER, JOSEF: The Effects of Framing, Reflection, Probability, and Payoff on Risk Preference in Choice Tasks, *Organizational Behavior and Human Decision Processes* 1999, 204-231
- KUHN, ANDRÉ /VILLET AZ, PATRICE /JOYET, ALINE/WILLI, FLORIAN: Öffentliche Meinung und Strenge der Richter, *Crimscope* Nr. 19, Lausanne 2000
- KULKA, RICHARD A./KESSLER, JOAN B.: Is Justice Really Blind? – The Influence of Liti-gant Physical Attractiveness on Juridical Judgment, *Journal of Applied Social Psy-chology* 1978, 366-381
- KUMMER, MAX: Grundriss des Zivilprozessrechts, 4. Aufl., Bern 1984

- KUNZ, KARL-LUDWIG/BÄCHTOLD, ANDREA/KIENER, REGINA/HEINE, GÜNTER/TSCHANNEN, PIERRE/KÄLIN, WALTER/KINZIG, JÖRG: Universität Bern zum Vorentwurf der Arbeitsgruppe „Verwahrung“, Jusletter vom 13. Dezember 2004, Rz. 7 ff., erhältlich unter [www.cx.unibe.ch/krim/Vernehmlassung% 20Uni.pdf](http://www.cx.unibe.ch/krim/Vernehmlassung%20Uni.pdf) (besucht am 15. Februar 2005)
- KUPFER SCHNEIDER, ANDREA: Shattering Negotiation Myths: Empirical Evidence on the Effectiveness of Negotiation Style, *Harvard Negotiation Law Review* 2002, 143-233
- KURAN, TIMUR/SUNSTEIN, CASS R.: Controlling Availability Cascades, *Stanford Law Review* 1999, 683 ff.; Nachdruck in: SUNSTEIN (Hrsg.), 374-397
- LABINE, SUSAN J./LABINE, GARY: Determination of Negligence and the Hindsight Bias, *Law and Human Behavior* 1996, 501-516
- LAIBSON, DAVID/ZECKHAUSER, RICHARD: Amos Tversky and the Ascent of Behavioral Economics, *Journal of Risk and Uncertainty* 1998, 7-47
- LANCE, CHARLES E./LAPOINTE, JULIE A./STEWART, AMY M.: A Test of the Context Dependency of Three Causal Models of Halo Rater Error, *Journal of Applied Psychology* 1994, 332-340
- LANDES, ELISABETH M./POSNER, RICHARD A.: The Economics of the Baby Shortage, *Journal of Legal Studies* 1978, 323-348
- LANDY, DAVID/SIGALL, HAROLD: Beauty Is Talent: Task Evaluation as a Function of the Performer's Physical Attractiveness, *Journal of Personality & Social Psychology* 1974, 299-304
- LANGEVOORT, DONALD C.: Organized Illusions: A Behavioral Theory of Why Corporations Mislead Stock Market Investors (and Cause Other Social Harms), *University of Pennsylvania Law Review* 1997, 101-172
- Behavioral Theories of Judgement and Decision Making in Legal Scholarship: A Literature Review, *Vanderbilt Law Review* 1998, 1499-1540
- LAPLACE, PIERRE-SIMON: Philosophischer Versuch über die Wahrscheinlichkeit, Leipzig 1932
- LARNTZ, KINLEY: Reanalysis of Vidmar's Data on the Effects of Decision Alternatives on Verdicts of Simulated Jurors, *Journal of Personality and Social Psychology* 1975, 123-140
- LEE, JU-WHEI/YATES, J.FRANK/SHINOTSUKA, HIROMI/SINGH, RAMADHAR/ ONGLATCO, MARY LOU UY/YEN, NAI-SHENG/GUPTA, MEENAKSHI/BHATNAGAR, DEEPTI: Cross-National Differences in Overconfidence, *Asian Journal of Psychology* 1995, 63-69
- LEFF, ARTHUR A.: Economic Analysis of Law: Some Realism About Nominalism, *Virginia Law Review* 1974, 451-482
- LEITER, BRIAN: Holmes, Economics, and Classical Realism, in: BURTON, STEVEN J. (Hrsg.): *The Path of the Law and Its Influence: The Legacy of Oliver Wendell Holmes, Jr*, Cambridge 2000, 285-325
- American Legal Realism, in: EDMUNDSON, WILLIAM/GOLDING, MARTIN (Hrsg.): *The Blackwell Guide to Philosophy of Law and Legal Theory*, Oxford 2004, 50-66
- LEMPERT, RICHARD: Modeling Relevance, *Michigan Law Review* 1977, 1021-1057

- The New Evidence Scholarship: Analyzing the Process of Proof, Boston University Law Review 1986, 439-477
- Juries, Hindsight, And Punitive Damage Awards: Failures of a Social Science Case for Change, DePaul Law Review 1999, 867-893
- LEVENTHAL, GLORIA/KRATE, RONALD: Physical Attractiveness and Severity of Sentencing, Psychological Reports 1977, 315-318
- LEVIN, IRWIN P./SCHNEIDER, SANDRA L./GAETH, GARY J.: All Frames Are not Created Equal: A Typology and Critical Analysis of Framing Effects, Organizational Behavior and Human Decision Processes 1998, 149-188
- LEWIS, CHARLES/KEREN, GIDEON: On the Difficulties Underlying Bayesian Reasoning: A Comment on Gigerenzer and Hoffrage, Psychological Review 1999, 411-416
- LICHTENSTEIN, SARAH/FISCHHOFF, BARUCH: Do Those Who Know More Also Know More about How Much They Know?, Organizational Behavior and Human Performance 1977, 159-183
- /- /PHILLIPS, LAWRENCE D.: Calibration of Probabilities: The State of the Art to 1980, in: KAHNEMAN/SLOVIC/TVERSKY (Hrsg.), 306-334
- LINDLEY, DENNIS: Foundations, in: AYTON/WRIGHT (Hrsg.), 3-15
- LINDSEY, SAMUEL/HERTWIG, RALPH/GIGERENZER, GERD: Communicating Statistical DNA Evidence, Jurimetrics 2003, 147-163
- LIPSHITZ, RAANAN/BARAK, DALYA: Hindsight Wisdom: Outcome Knowledge and the Evaluation of Decisions, Acta Psychologica 1995, 105-125
- LLEWELLYN, KARL N.: The Bramble Bush, New York 1930
- LLEWELLYN, KARL N.: Some Realism about Realism – Responding to Dean Pound, Harvard Law Review 1931, 1222-1264
- LOEWENSTEIN, GEORGE/ISSACHAROFF, SAMUEL/CAMERER, COLIN/BABCOCK, LINDA: Self-Serving Assessments of Fairness and Pretrial Bargaining, Journal of Legal Studies 1993, 135-159
- LOFTUS, ELIZABETH F.: To File, Perchance To Cheat, Psychology Today, April 1985, 34 ff.
- /WAGENAAR, WILLEM A. : Lawyers' Predictions of Success, Jurimetrics 1987, 437-453
- LOPES, LOLA L.: The Rhetoric of Irrationality, Theory & Psychology 1991, 65-82
- LORD, CHARLES G./ROSS, LEE/LEPPER, MARK R.: Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence, Journal of Personality and Social Psychology 1979, 2098-2109
- LÖSCHPER, GABRIELE: Bausteine für eine psychologische Theorie richterlichen Urteilens, Baden-Baden 1999
- LYON, DON/SLOVIC, PAUL: Dominance of Accuracy Information and Neglect of Base-Rates in Probability Estimates, Acta Psychologica 1976, 287-298
- LYON, THOMAS D./KOEHLER, JONATHAN J.: The Relevance Ratio: Evaluating the Probative Value of Expert Testimony in Child Sexual Abuse Cases, Cornell Law Review 1996, 43-78
- MAASSEN, BERNHARD U.: Beweismassprobleme im Schadenersatzprozess, Köln etc. 1975

Allgemeiner Teil

- MACCOUN, ROBERT: Why a Psychologist Won the Nobel Prize in Economics, American Psychological Society Observer 2002, 1-8
- MALOUFF, JOHN/SCHUTTE, NICOLA S.: Shaping Juror Attitudes: Effects of Requesting Different Damage Amounts in Personal Injury Trials, Journal of Social Psychology 491-497
- MALSCH, MARIJKE: Lawyers' Predictions of Judicial Decisions, Diss. Leiden 1989
- MARK, MELVIN M./REITER BOBURKA, RENEE/EYSELL, KRISTEN M./COHEN, LAURIE L./MELLOR, STEVEN: „I couldn't have seen it coming“: The Impact of Negative Self-Relevant Outcomes on Retrospections About Foreseeability, Memory 2003, 443-454
- MARKOWITZ, HARRY: The Utility of Wealth, Journal of Political Economy 1952, 151-158; erhältlich unter <http://cowles.econ.yale.edu/P/cp/p00b/p0057.pdf> (besucht am 23. März 2005)
- MCCAFFREY, EDWARD J.: Cognitive Theory and Tax, UCLA Law Review 1994, 1961-1981; Nachdruck in: SUNSTEIN (Hrsg.), 398-421. Zitiert wird der Nachdruck.
- McFATTER, ROBERT: Sentencing Strategies and Justice: Effects of Punishment Philosophy On Sentencing Decisions, Journal of Personality and Social Psychology 1978, 1490-1500
- MEEHL, PAUL A.: Clinical versus Statistical Prediction, Minneapolis 1954
- MELLERS, BARBARA/HERTWIG, RALPH/KAHNEMAN, DANIEL: Do Frequency Representations Eliminate Conjunction Effects?, Psychological Science 2001, 269-275
- MISES, RICHARD VON: Wahrscheinlichkeit, Statistik und Wahrheit, 4. Aufl., Wien/New York 1972
- MITCHELL, GREGORY A.: Why Law and Economics' Perfect Rationality Should Not Be Traded for Behavioral Law and Economics' Equal Incompetence, Georgetown Law Journal 2002, 67-167
- Taking Behavioralism Too Seriously? The Unwarranted Pessimism of the New Behavioral Analysis of Law, William & Mary Law Review 2002, 1907-2021
- MURPHY, KEVIN R./JAKO, ROBERT A./ANHALT, REBECCA L.: Nature and Consequences of Halo Error: A Critical Analysis, Journal of Applied Psychology 1993, 218-225
- MURPHY, MARCIA J./HELKAMP, DAVID T.: Attractiveness and Personality Warmth: Evaluations of Paintings Rated by College Men and Women, Perceptual & Motor Skills 1976, 1163-1166
- MUSCH, JOCHEN: Personality Differences in Hindsight Bias, Memory 2003, 473-489
- MUSIELAK, HANS JOACHIM: Die Grundlagen der Beweislast im Zivilprozessrecht, Berlin/New York 1978
- MUSIELAK, HANS-JOACHIM: Das Överviktsprincip: Zum Verhältnis von richterlicher Überzeugung und Wahrscheinlichkeit, in: LÜDERITZ, ALEXANDER/SCHRÖDER, JOCHEN: Festschrift für Gerhard Kegel, Frankfurt 1977, 451-471
- MÜSSELER JOCHEN/PRINZ, WOLFGANG: Allgemeine Psychologie, Heidelberg/Berlin 2002

- MUSSWEILER, THOMAS/STRACK, FRITZ: Explaining the Enigmatic Anchoring Effect: Mechanisms of Selective Accessibility, *Journal of Personality and Social Psychology* 1997, 437-446
- /- /PFEIFFER, TIM: Overcoming the Inevitable Anchoring Effect: Considering the Opposite Compensates for Selective Accessibility, *Personality and Social Psychology Bulletin* 2000, 1142-1150
- /- Numeric Judgments under Uncertainty: The Role of Knowledge in Anchoring, *Journal of Experimental Social Psychology* 2000, 495-518
- The Durability of Anchoring Effects, *European Journal of Social Psychology* 2001, 431-442
- /- Considering the Impossible: Explaining the Effects of Implausible Anchors, *Social Cognition* 2001, 145-160
- MYNATT, CLIFFORD R./DOHERTY, MICHAEL E./DRAGAN, WILLIAM: Information Relevance, Working Memory, and the Consideration of Alternatives, *The Quarterly Journal of Experimental Psychology* 1993, 759-788
- NACK, ARMIN: Beweiswürdigung beim Indizienbeweis, *Kriminalistik* 1995, 466-470
- NELL, ERNST LUDWIG: Wahrscheinlichkeitsurteile in juristischen Entscheidungen, Diss. Bayreuth 1982, Berlin 1983
- NESSON, CHARLES: Reasonable Doubt and the Permissive Inference: The Value of Complexity, *Harvard Law Review* 1979, 1187-1215
- NICKERSON, RAYMOND S.: Confirmation Bias: A Ubiquitous Phenomenon in Many Guises, *Review of General Psychology* 1998, 175-220
- NISBETT, RICHARD/ROSS, LEE: Human Inference: Strategies and Shortcomings of Social Judgment, Englewood Cliffs 1980
- NOLL, ROGER C./KRIER, JAMES E.: Some Implications of Cognitive Psychology for Risk Regulation, *Journal of Legal Studies* 1990, 747-779
- NONN, MICHAEL: Die Beweiswürdigung im Zivilprozess unter besonderer Berücksichtigung des Kantons Basel-Landschaft, Basel 1996
- NORTHCROFT, GREGORY B./NEALE, MARGARET A.: Experts, Amateurs, and Real Estate: An Anchoring-and-Adjustment Perspective on Property Pricing Decisions, *Organizational Behavior & Human Decision Processes* 1987, 84-97
- NOTZ, WILLIAM W./BOSCHMANN, IRVIN/BRUNING, NEALIA S.: Punishment Without Cause: Regression and the Effects of Leader Attribution Error, *Journal of Applied Social Psychology* 2001, 2401-2416
- OFIR, CHEZY/MAZURSKY, DAVID: Does a Surprising Outcome Reinforce or Reverse the Hindsight Bias?, *Organizational Behavior and Human Decision Making Processes* 1997, 51-57
- OFTINGER, KARL/STARK, EMIL W.: Schweizerisches Haftpflichtrecht, Erster Band: Allgemeiner Teil, 5. Aufl., Zürich 1995
- OLIPHANT, HERMAN: A Return to Stare Decisis, *American Bar Association Journal* 1928, 71-76, 107, 159-162, 75

Allgemeiner Teil

- OSKAMP, STUART: Overconfidence in Case-Study Judgments, *Journal of Consulting Psychology* 1965, 261-265; Nachdruck in: KAHNEMAN/SLOVIC/ TVERSKY (Hrsg.), 287-293. Zitiert wird der Nachdruck.
- OSWALD, MARGRIT: Urteile über den Repräsentativitätsheurismus, *Archiv für Psychologie* 1986, 113-125
- Hypothesentesten: Suche und Verarbeitung hypothesenkonformer und hypothesenkonträrer Informationen, in: HELL/FIEDLER/GIGERENZER (Hrsg.), 189-212
 - Psychologie richterlichen Strafens, Stuttgart 1994
- PENNINGTON, DONALD C.: Being Wise After the Event: An Investigation of Hindsight Bias, *Current Psychological Research* 1981, 271-282
- Witnesses and Their Testimony: Effects of Ordering on Juror Verdicts, *Journal of Applied Social Psychology* 1982, 318-333
 - /LLOYD-BOSTOCK, SALLY (Hrsg.): *The Psychology of Sentencing*, Oxford 1987
- PERFORS, AMY: What's in a Name? – The effect of sound symbolism on perception of facial attractiveness, 2004; erhältlich unter www.mit.edu/~perfors/hotornot.pdf (besucht am 5. Februar 2005)
- PETERS, DOROTHEE: Richter im Dienste der Macht: Zur gesellschaftlichen Verteilung der Kriminalität, Stuttgart 1973
- PETERS, PHILIP G.: Hindsight Bias and Tort Liability: Avoiding Premature Conclusions, *Arizona State Law Review* 1999, 1277-1314
- PETERSON, JOSEPH L./GAENSSLEN, R. E.: Developing Criteria for Model External DNA Proficiency Testing: Final Report National Institute of Justice, 2001
- PEZZO, MARK V.: Surprise, Defence or Making Sense: What Removes Hindsight Bias?, *Memory* 2003, 421-441
- PHILLIPS, LAWRENCE D./EDWARDS, WARD: Conservatism in Simple Probability Inference Tasks, *Journal of Experimental Psychology* 1966, 346-357
- PIEHL, JOCHEN: Integration of Information in the "Courts": Influence of Physical Attractiveness on Amount of Punishment for a Traffic Offender, *Psychological Reports* 1977, 551-556
- PLOUS, SCOTT: *The Psychology of Judgment and Decision Making*, New York 1993
- POHL, RÜDIGER F./HELL, WOLFGANG: No Reduction in Hindsight Bias after Complete Information and Repeated Testing, *Organizational Behaviour and Human Decision Processes* 1996, 49-58
- /EISENHAEUER, MARKUS/HARDT, OLIVER: SARA: A cognitive process model to simulate the anchoring effect and hindsight bias, *Memory* 2003, 337-356
 - (Hrsg.): *Cognitive Illusions*, Hove 2004
- POLINSKY, A. MITCHELL: Economic Analysis as a Potentially Defective Product: A Buyer's Guide to Posner's Economic Analysis of Law, *Harvard Law Review* 1974, 1655-1681
- POOLE, DEBRA/LINDSAY, D. STEPHEN: Assessing the Accuracy of Young Children's Reports: Lessons from the Investigation of Child Sexual Abuse, *Applied & Preventive Psychology* 1998, 1-26

- POSNER, RICHARD A.: The Economic Approach to Law, Texas Law Review 1975, 757-787
- Wealth Maximization Revisited, Notre Dame Journal of Law, Ethics & Public Policy 1985, 85-105
 - The Regulation of the Market in Adoption, Boston University Law Review 1987, 59-72
 - The Future of Law and Economics: A Comment on Ellickson, Chicago-Kent Law Review 1989, 57-70
 - What Do Justices Maximize? (The Same Thing As Everybody Else Does), Supreme Court Economic Review 1993
 - Sex and Reason, Cambridge 1994 (Paperback)
 - Rational Choice, Behavioral Economics, and the Law, Stanford Law Review 1998, 1551-1575
 - Economic Analysis of Law, 6. Aufl., New York 2002
- POUND, ROSCOE: Enforcement of Law, Green Bag 1908, 401-415
- PRENTICE, ROBERT A.: Chicago Man, K-T Man, And the Future of Behavioral Law and Economics, Vanderbilt Law Review 2003, 1663-1777
- PRIEST, GEORGE L./KLEIN, BENJAMIN: The Selection of Disputes for Litigation, Journal of Legal Studies 1984, 1-55
- PRONIN, EMILY/LIN, DANIEL Y./ROSS, LEE: The Bias Blind Spot: Perceptions of Bias in Self Versus Others, Personality and Social Psychology Bulletin 2002, 369-381
- PYSZCZYNSKI, TOM/GREENBERG, JEFF: Toward an Integration of Cognitive and Motivational Perspectives on Social Inference: A Biased Hypothesis Testing Model, in: BERKOWITZ, LEONARD (Hrsg.): Advances in Experimental Social Psychology, Vol. 30, San Diego etc. 1987, 297-340
- QUALLS, WILLIAM J./PUTO, CHRISTOPHER P.: Organizational Climate and Decision Framing: An Integrated Approach to Analyzing Industrial Buying Decisions, Journal of Marketing Research 1989, 179-192
- RACHLINSKI, JEFFREY J.: Gains, Losses and the Psychology of Litigation, Southern California Law Review 1996, 113-185
- The Uncertain Psychological Case for Paternalism, Northwestern University Law Review 1997, 1165-1225
 - A Positive Psychological Theory of Judging in Hindsight, University of Chicago Law Review 1998, 571-625
 - Heuristics and Biases in the Courts: Ignorance or Adaptation?, Oregon Law Review 2000, 61-102
 - The „New“ Law and Psychology: A Reply to Critics, Sceptics and Cautious Supporters, Cornell Law Review 2000, 739-766, 740.
- RAITZ ALLAN/GREENE, EDITH/GOODMAN, JANE /LOFTUS, ELIZABETH F.: Determining Damages: The Influence of Expert Testimony on Jurors' Decision Making, Law and Human Behavior 1990, 385-395

Allgemeiner Teil

- RAMSEY, FRANK P.: Truth and Probability, London 1926, in: BRAITHWAITE, RICHARD B. (Hrsg.), The Foundations of Mathematics and other Logical Essays, New York 1931, 156-198
- REDDING, RICHARD E./REPPUCH, N. DICKSON: Effects of Lawyers' Socio-political Attitudes on Their Judgments of Social Science in Legal Decision Making, Law and Human Behavior 1999, 31-54
- REDONDO, SANTIAGO/GARRIDO, VICENTE /PÉREZ, JORGE / BARBERET, ROSEMARY: Advances in Psychology and Law, Berlin etc. 1997, 215-226
- REHBERG, JÖRG: Strafrecht I, 5. Aufl. Zürich 1993
- REHBINDER, MANFRED: Karl N. Llewellyn als Rechtssoziologe, Kölner Zeitschrift für Soziologie und Sozialpsychologie 1966, 532-556
- Neues über Leben und Werk von Eugen Ehrlich, in: KAULBACH, FRIEDRICH/KRAWIETZ, WERNER (Hrsg.): Festschrift für Helmut Schelsky zum 65. Geburtstag, Berlin 1978, 403-418
 - Einleitung und Kommentar zu den Art. 319-330a OR, Bern 1985
 - Rechtssoziologie, 5. Aufl. München 2003.
- RITOV, ILANA/BARON, JONATHAN: Status-quo and Omission NBias, Journal of Risk and Uncertainty 1992, 49-61
- /- Outcome Knowledge, Regret, and Omission Bias, Organizational Behavior and Human Decision Processes 1995, 119-127
 - Anchoring in Simulated Competitive Market Negotiation, Organizational Behavior and Human Decision Processes 1996, 16-25
- ROBBEN, HENRY S. J./WEBLEY, PAUL/WEIGEL, RUSSELL H./WARNERYD, KARL-ERIK/KINSEY, KARYL A./HESSING, DICK J./ALVIRA MARTIN, FRANCISCO/ELFFERS, HENK/WAHLUND, RICHARD/VAN LANGENHOVE, LUK: Decision frame and opportunity as determinants of tax cheating: An international experimental study, Journal of Economic Psychology 1990, 341-364
- ROBBENNOLT, JENNIFER K./STUDEBAKER, CHRISTINA A.: Anchoring in the Courtroom: The Effects of Caps on Punitive Damages, Law and Human Behavior 1999, 353-373
- ROBERTO, VITO: Schweizerisches Haftpflichtrecht, Zürich 2002
- ROBINS, RICHARD W./CRAIK, KENNETH: Is There A Citation Bias in the Judgement and Decision Literature?, Organizational Behavior and Human Decision Processes 1993, 225-244
- ROESE, NEAL J./OLSON, JAMES M.: What Might Have Been: The Social Role of Counterfactual Thinking, New York 1995
- ROSS, LEE/LEPPER, MARK R./HUBBARD, MICHAEL: Perseverance in Self Perception and Social Perception: Biased Attributional Processes in the Debriefing Paradigm, Journal of Personality and Social Psychology 1975, 880-892
- /ANDERSON, CRAIG A.: Shortcomings in the Attribution Process, in: TVERSKY/KAHNEMAN/SLOVIC (Hrsg.), 129-152
 - /SHESTOWSKY, DONN: Contemporary Psychology's Challenges to Legal Theory and Legal Practice, Northwestern University Law Review 2003, 1081-1114

- ROSS, MICHAEL/SICOLY, FIORE: Egocentric Biases in Availability and Attribution, in: KAHNEMAN/SLOVIC/TVERSKY (Hrsg.), 179-189
- ROSZKOWSKI, MICHAEL J./SNELBECKER, GLEN E.: Effects of "Framing" on Measures of Risk Tolerance: Financial Planners are Not Immune, *Journal of Behavioral Economics* 1990, 237-246
- ROYZMAN, EDWARD/BARON, JONATHAN: The Preference for Indirect Harm, *Social Justice Research* 2002, 165-184
- RUSSO, J. EDWARD/SHOEMAKER, PAUL J.H.: Managing Overconfidence, *Sloan Management Review* 1992, 7-17
- SAKS, MICHAEL J./KIDD, ROBERT F.: Human Information Processing and Adjudication: Trial by Heuristics, *Law & Society Review* 1980, 123-160
- SALES, BRUCE DENNIS (Hrsg.), *The Trial Process*, New York 1981, 481-498
- SAMUELSON, PAUL A.: Problems of Methodology – Discussion, *American Economic Review Papers and Proceedings* 1963, 231-236
- SAMUELSON, WILLIAM/ZECKHAUSER, RICHARD: Status Quo Bias in Decision Making, *Journal of Risk and Uncertainty* 1988, 7-59
- SANNA, LAWRENCE J./SCHWARZ, NORBERT/STOCKER, SHEVAUN L.: When Debiasing Backfires: Accessible Content and Accessibility Experiences in Debiasing Hindsight, *Journal of Experimental Psychology* 2002, 497-502
- SAVAGE, LEONHARD: *The Foundations of Statistics*, New York 1954
- SCHLEGEL, JOHN H.: *American Legal Realism and Empirical Social Science*, Durham 1995
- SCHMID, NIKLAUS: *Strafprozessrecht*, 2. Aufl., Zürich 1993
- SCHMITT, MANFRED: Schönheit und Talent: Untersuchungen zum Verschwinden des Halo-Effekts, *Zeitschrift für experimentelle und angewandte Psychologie* 1992, 475-492
- SCHNEIDER, EGON: *Logik für Juristen*, 5. Aufl. München 1999
- SCHNELL, RAINER/HILL, PAUL B./ESSER, ELKE: *Methoden der empirischen Sozialforschung*, 7. Aufl. München 2005
- SCHREIBER, RUPERT: *Theorie des Beweiswerts für Beweismittel im Zivilprozessrecht*, Berlin etc. 1968
- SCHULZ-HARDT, STEFAN/FREY, DIETER/FAGO, KRISTIANE/KICI, GÜLER: Selektive Informationssuche und Gruppenheterogenität: Der Einfluss verschiedener Formen der Gruppenheterogenität auf Selbstbestätigungsprozesse bei Entscheidungen, *Gruppendynamik* 1999, 161-174
- SCHULTZ, HANS: *Rechtsprechung und Praxis zum Strassenverkehrsrecht in den Jahren 1983-1987*, Bern 1990, 94 f.
- SCHUM, DAVID A./MARTIN, ANNE W.: Formal and Empirical Research on Cascaded Inference, *Law and Society Review* 1982, 105-151; Nachdruck in: HASTIE, REID (Hrsg.): *Inside the Juror*, Cambridge 1993, 136-174. Zitiert wird der Nachdruck.
- SCHWEITZER, MAURICE: Disentangling Status Quo and Omission Effects: An Experimental Analysis, *Organizational Behavior and Human Decision Making Processes* 2002, 457-476

- SCHWEIZER, MARK: The Offer/Asking Price Gap and Term of Possession – an Experiment, unpublizierte Master-These Universität Michigan, 2002; erhältlich unter www.decisions.ch/publikationen.html (besucht am 1. Mai 2005)
- SEDLMEIER, PETER/GIGERENZER, GERD: Teaching Bayesian Reasoning in Less Than Two Hours, *Journal of Experimental Psychology: General* 2001, 380-400
- SEGAL, JEFFREY A./SPAETH, HAROLD J.: Supreme Court and the Attitudinal Model Revisited, Cambridge 2002
- SEIDMAN DIAMOND, SHARI: The Challenges of Socio-Legal Research on Decision Making: Psychological Successes and Failures, *Journal of Law and Society* 1995, 78-84
- SHAFIR, ELDAR: Cognitive Psychology of Decision Biases, in: Smelser, Neil J./Baltes, Paul B.: *International Encyclopedia of the Social & Behavioral Sciences*, Amsterdam etc. 2001, 3296-3300
- SHAFIR, SHARONI/WAITE, TOM A./SMITH, BRIAN H.: Context-Dependent Violations of Rational Choice in Honeybees (*Apis Mellifera*) and Gray Jaws (*Perisoreus Canadensis*), *Behavioral Ecology and Sociobiology* 2002, 180-187
- SHAPIRO, FRED R.: The Most Cited Legal Scholars, *Journal of Legal Studies* 2000, 409-426
- The Most-Cited Legal Books Published Since 1978, *Journal of Legal Studies* 2000, 397-406
- SHAPIRO, STEWART/SPENCE, MARK T.: Mind Over Matter? The Inability to Counteract Contrast Effects Despite Conscious Effort, *Psychology & Marketing* 2005, 225-245
- SIGALL, HAROLD/OSTROVE, NANCY: Beautiful but Dangerous: Effects of Offender Attractiveness and Nature of Crime on Juridic Judgment, *Journal of Personality and Social Psychology* 1975, 410-414
- SIMON, HERBERT A.: A Behavioral Model of Rational Choice, *Quarterly Journal of Economics* 1955, 99-118
- Models of Man: Social and Rational, New York 1957
 - Rational Decision-Making in Business Organizations, *Economic Sciences* 1978, 343-371, 356; erhältlich unter <http://nobelprize.org/economics/laureates/1978/simon-lecture.pdf> (besucht am 7. April 2005)
 - Bounded Rationality, in: EATWELL, JOHN/MILGATE, MURRAY/NEWMAN, PETER: *The New Palgrave: a dictionary of economics*, London 1987, Bd. 1, 266-267
- SIMON, RITA JAMES: „Beyond Reasonable Doubt“ – An Experimental Attempt at Quantification, *Journal of Applied Behavioral Science* 1970, 203-209
- /MAHAN, LINDA: Quantifying Burdens of Proof: A View from the Bench, the Jury and the Classroom, *Law and Society Review* 1971, 319-330
- SIMONSON, ITAMAR: Choice Based on Reason: The Case of Attraction and Compromise Effects, *Journal of Consumer Research* 1989, 158-174
- /TVERSKY, AMOS: Choice in Context: Tradeoff Contrast and Extremeness Aversion, *Journal of Marketing Research* 1992, 281-295

- SINE, WESLEY DAVID/SHANE, SCOTT/DI GREGORIO, DANTE: The Halo Effect and Technology Licensing: The Influence of Institutional Prestige on the Licensing of University Inventions, *Management Science* 2003, 478-496
- SINGER, JOSEPH WILLIAM: Legal Realism Now, *California Law Review* 1988, 467-544
- SINGER, PETER: *Praktische Ethik*, Stuttgart 1984 (deutsche Übersetzung der englischen Originalausgabe von 1979)
- SISK, GREGORY C. /HEISE, MICHAEL/MORRIS, ANDREW P.: Charting the Influences on the Judicial Mind: An Empirical Study of Judicial Reasoning, *New York University Law Review* 1998
- SLOMAN, STEVEN A.: Two Systems of Reasoning, in: GILOVICH/GRIFFIN/KAHNEMAN (Hrsg.), 379-396
- SLOVIC, PAUL/FISCHHOFF, BARUCH: On the Psychology of Experimental Surprises, *Journal of Experimental Psychology: Human Perception and Performance* 1977, 544-551
- SMITH, EDWARD D./HED, ANITA: EFFECTS of Offender Age and Attractiveness on Sentencing by Mock Jurors, *Psychological Reports* 1979, 691-694
- SMITH, KENNETH W./KINSEY, KARYL A.: Understanding Taxpaying Behavior: A Conceptual Framework with Implications for Research, *Law and Society Review* 1987, 639-663
- SORENSEN, TEENA/SNOW, BARBARA: How Children Tell: The Process of Disclosure in Child Sexual Abuse, *Child Welfare* 1991, 3-15
- SPRANCA, MARK/MINSK, ELISA/BARON, JONATHAN: Omission and Commission in Judgment and Choice, *Journal of Experimental Social Psychology* 1991, 76-105
- STANOVICH, KEITH E./WEST, RICHARD F.: Individual Differences in Reasoning: Implications for the Rationality Debate, *Behavioral and Brain Sciences* 2000, 645-726
- STEPHAN, COOKIE/TULLY, JUDY CORDER: The Influence of Physical Attractiveness of a Plaintiff on the Decisions of Simulated Jurors, *Journal of Social Psychology* 1977, 149-150
- STEWART, IAN: Der Trugschluss des Ermittlers, *Spektrum der Wissenschaft* Juli 1997, 8.
- STEWART, JOHN E.: Defendant's Attractiveness as a Factor in the Outcome of Criminal Trials: An Observational Study, *Journal of Applied Social Psychology* 1980, 348-361
- STIGLER, STEPHEN M.: Who discovered Bayes's Theorem?, *The American Statistician* 1983, 290-296
- STRATENWERTH, GÜNTER: *Schweizerisches Strafrecht, Allgemeiner Teil I: Die Straftat*, 2. Aufl., Bern 1996
- /JENNY, GUIDO: *Schweizerisches Strafrecht, Besonderer Teil I: Straftaten gegen Individualinteressen*, 6. Aufl., Bern 2003
- SUMMERS, ROBERT S.: *Instrumentalism and American Legal Theory*, Ithaca 1982
- SUNSTEIN, CASS R.: Behavioral Analysis of Law, *University of Chicago Law Review* 1997, 1175-1195
- /KAHNEMAN, DANIEL/SCHADKE, DAVID: Assessing Punitive Damages (With Notes on Cognition and Valuation in Law), *Yale Law Journal* 1998, 2071-2152

Allgemeiner Teil

- (Hrsg.): Behavioral Law and Economics, Cambridge 2000
- Switching the Default Rule, New York University Law Review 2002, 106-134
- /THALER, RICHARD H.: Libertarian Paternalism Is Not an Oxymoron, University of Chicago Law Review 2003, 1159-1202
- Moral Heuristics and Moral Framing, Minnesota Law Review 2004, 1556-1597
- SVENSON, OLA: Are We All Less Risky and More Skillful Than Our Fellow Drivers?, Acta Psychologica 1980, 143-148
- TARONI, FRANCO/AITKEN, COLIN: Probabilités et preuve par l'ADN dans les affaires civiles et criminelles: Questions de la cour et réponses fallacieuses des experts, Schweizerische Zeitschrift für Strafrecht 1998, 291-313
- /MANGIN, PATRICE/BÄR, WALTER: Die Interpretation des Beweiswertes von DNA-Untersuchungen in sachverständigen Gutachten, Schweizerische Zeitschrift für Strafrecht 1999, 439-445
- TAYLOR, SHELLEY E./BROWN, JONATHAN D.: Illusion and Well-Being: A Social Psychological Perspective on Mental Health, Psychological Bulletin 1988, 193-210
- Positive Illusions: Creative Self-Deception and the Healthy Mind, New York 1989, 214
- THALER, RICHARD H.: Toward a Positive Theory of Consumer Choice, Journal of Economic Behavior and Organization 1980, 39-60
- THIBAUT, JOHN/WALKER, LAURENS: Procedural Justice: A Psychological Analysis, Hillsdale 1975
- THOMPSON, WILLIAM C./SCHUMANN, EDWARD L.: Interpretation of Statistical Evidence in Criminal Trials: The Prosecutor's Fallacy and the Defense Attorney's Fallacy, Law and Human Behavior 1987, 167-187
- /TARONI, FRANCO/AITKEN, COLIN C.G.: How the Probability of a False Positive Affects the Value of DNA Evidence, Journal of Forensic Science 2003, 47-54
- THORNDIKE, EDWARD L.: A Constant Error in Psychological Ratings, Journal of Applied Psychology 1920, 25-29
- TORNGREN, GUSTAF/MONTGOMERY, HENRY: Worse Than Chance? Performance and Confidence Among Professionals and Laypeople in the Stock Market, Journal of Behavioral Finance 2004, 148-153
- TRECHSEL, STEFAN: Schweizerisches Strafgesetzbuch: Kurzkommentar, Zürich 1989
- /NOLL, PETER: Schweizerisches Strafrecht, Allgemeiner Teil I, 6. Aufl., Zürich 2004
- TRIBE, LAWRENCE: Trial by Mathematics: Precision and Ritual in the Legal Process, Harvard Law Review 1971, 1329-1351
- TVERSKY, AMOS/KAHNEMAN, DANIEL: Availability: A heuristic for judging frequency and probability, Cognitive Psychology 1973, 207-232
- /- Judgment under Uncertainty: Heuristics and Biases, Science 1974, 1124-1131
- /- Causal schemas in judgments under uncertainty, in: FISHBEIN, MARTIN (Hrsg.): Progress in social psychology, Hillsdale 1980, 49-72

- /- The Framing of Decisions and the Psychology of Choice, Science 1981, 453-458
 - /- Judgments of and by Representativeness, in: KAHNEMAN/SLOVIC/TVERSKY (Hrsg.), 84-100
 - /- Evidential Impact of Base Rates, in: KAHNEMAN/SLOVIC/TVERSKY (Hrsg.), 153-160
 - /- Rational Choice and the Framing of Decisions, Journal of Business 1986, 251-278
 - /- Loss Aversion in Riskless Choice: A Reference-Dependent Model, Quarterly Journal of Economics 1991, 1039-1061
 - /- Advances in Prospect Theory: Cumulative Representation of Uncertainty, Journal of Risk and Uncertainty 1992, 297-323; Nachdruck in: KAHNEMAN/TVERSKY (Hrsg.), 44-65. Zitiert wird der Nachdruck.
 - /SIMONSON, ITAMAR: Context-Dependent Preferences, Management Science 1993, 117-185; gekürzter Nachdruck in: KAHNEMAN/TVERSKY (Hrsg.), 518-527. Zitiert wird der Nachdruck.
- URY, WILLIAM: Getting Past No, New York 1993
- VAN DUYNE, PETRUS: Simple Decision Making, in: PENNINGTON, DONALD C./LLOYD-BOSTOCK, SALLY (Hrsg.): The Psychology of Sentencing, Oxford 1987, 143-156
- VAN POUCKE, DIRK/BUELENS, MARC: Predicting the Outcome of a Two-Party Price Negotiation: Contribution of Reservation Price, Aspiration Price and Opening Offer, Journal of Economic Psychology 2002, 67-76
- VAUCHER, STEVE/STORZ, RENATE/RÔNEZ, SIMONE: Strassenverkehrsdelinquenz und Rückfall, Neuchâtel 2000
- VERKERKE, J. HOULT: An Empirical Perspective on Indefinite Term Employment Contracts: Resolving the Just Cause Debate, Wisconsin Law Review 1995, 837-901
- VIDMAR, NEIL: Effects of Decision Alternatives on the Verdicts and Social Perception of Simulated Jurors, Journal of Personality and Social Psychology 1972, 211-218
- VILLEJOUBERT, GÄELLE/MANDEL, DAVID R.: The Inverse Fallacy: An Account of Deviations from Bayes's Theorem and the Additivity Principle, Memory & Cognition 2002, 171-178
- VOGEL, OSCAR/SPÜHLER, KARL: Grundriss des Zivilprozessrechts, 7. Aufl., Bern 2001
- VON HOLSTEIN, CARL-AXEL S. STAËL: Probabilistic Forecasting: An Experiment Related to Stock Market, Organizational Behavior and Human Performance 1972, 139-158
- VON NEUMANN, JOHN/MORGENSTERN, OSKAR: Theory of Games and Economic Behavior, 2. Auflage, Princeton 1947
- WAGENAAR, WILLEM A.: The Subjective Probability of Guilt, in: AYTON/WRIGHT (Hrsg.), 529-547
- WALLSTEN, THOMAS S. (Hrsg.): Cognitive Processes in Choice and Decision Behavior, Hillsdale 1980
- WASON, PETER C. : On the Failure to Eliminate Hypothesis in a Conceptual Task, Quarterly Journal of Experimental Psychology 1960, 129-140
- /JOHNSON-LAIRD, PHILIP N.: Psychology of Reasoning, London 1972

- WEIGEL, RUSSEL H./HESSING, DICK J./ELFFERS, HENK: Tax Evasion Research: A Critical Appraisal and Theoretical Model, *Journal of Economic Psychology* 1987, 215-235
- WEINSTEIN, NEIL D.: Unrealistic Optimism About Future Life Events, *Journal of Personality and Social Psychology* 1980, 806-820
- WEITEN, WAYNE: The Attraction-Leniency Effect in Jury Research: An Examination of External Validity, *Journal of Applied Social Psychology* 1980, 340-347
- WEITNAUER, HERMANN: Wahrscheinlichkeit und Tatsachenfeststellung, *Karlsruher Forum* 1966
- WELLS, FREDERIC L.: A Statistical Study of Literary Merit, *Archives of Psychology* 1907, 1-30
- WETZEL, CHRISTOPHER G./WILSON, TIMOTHY D./KORT, JAMES: The Halo Effect Revisited: Forewarned Is Not Forearmed, *Journal of Experimental Social Psychology* 1981, 427-439
- WHITE, SALLY B./NEALE, MARGARET A.: The Role of Negotiator Aspirations and Settlement Expectancies in Bargaining Outcomes, *Organizational Behavior and Human Decision Processes* 1994, 303-317
- WIDMER, PIERRE: Privatrechtliche Haftung, in: MÜNCH, PETER/GEISER, THOMAS (Hrsg.): Schaden – Haftung – Versicherung, Basel etc. 1999, 7-93
- WILSON, TIMOTHY D./HOUSTON, CHRISTOPHER E./ETLING, KATHRYN M./BREKKE, NANCY: A New Look at Anchoring Effects: Basic Anchoring and Its Antecedents, *Journal of Experimental Psychology: General* 1996, 387-402
- WRIGHT, WILLIAM/ANDERSON, URTON: Effects of Situation Familiarity and Financial Incentives on Use of the Anchoring and Adjustment Heuristic for Probability Assessment, *Organizational Behavior and Human Decision Processes* 1989, 68-82
- WUENSCH, KARL L./CASELOW, WILBUR A./MOORE, CHARLES H.: Effects of Defendant Attractiveness and Type of Crime on Juridic Judgment, *Journal of Social Behavior and Personality* 1991, 713-724
- YATES, J. FRANK/LEE, JU-WHEI/SIECK, WINSTON R./CHOI, INCHEOL/PRICE, PAUL C.: Probability Judgment across Cultures, in: GILOVICH/GRIFFIN/KAHNEMAN (Hrsg.), 271-291
- ZEBROWITZ, LESLIE A./MCDONALD, SUSAN M.: The Impact of Litigants' Baby-Facedness and Attractiveness on Adjudications in Small Claims Courts, *Law and Human Behavior* 1991, 603-623
- ZEISEL, HANS/KAYE, DAVID: Prove it with Figures: Empirical Methods in Law and Litigation, New York 1997
- ZIMMERMANN-ACKLIN, MARKUS: Euthanasie, 2. Aufl., Freiburg 2002

Quellenverzeichnis

- Arbeitsgruppe Sterbehilfe: Bericht an das Eidgenössische Justiz- und Polizeidepartement, März 1999; erhältlich unter [www.ofj.admin.ch/themen/stgb -sterbehilfe/b-bericht-d.pdf](http://www.ofj.admin.ch/themen/stgb/-sterbehilfe/b-bericht-d.pdf) (besucht am 3. Februar 2005)
- Arbeitsgruppe „Verwahrung“: Vorentwurf vom 15. Juli 2004 zur Änderung des Strafgesetzbuches in der Fassung vom 13. Dezember 2002 betreffend die Umsetzung von Artikel 123a BV (lebenslängliche Verwahrung) und einzelne nachträgliche Korrekturen am neuen Massnahmerecht; erhältlich unter www.ofj.admin.ch/themen/lebverwahr/vorentwurf-d.pdf (besucht am 15. Februar 2005)
- Beratungsstelle für Unfallverhütung: bfu aktuell 2004/5, Bern 2004; erhältlich unter www.bfu.ch/medien/bfu_aktuell/2004_05_d.pdf (besucht am 11. Februar 2005)
- Beratungsstelle für Unfallverhütung: Medienmitteilung vom 11. Februar 2005; erhältlich unter www.bfu.ch/medien/infomedia/index.htm (besucht am 17. Februar 2005)
- Botschaft zum Bundesgesetz über die Verwendung von DNA-Profilen im Strafverfahren und zur Identifizierung von unbekannten und vermissten Personen; erhältlich unter www.admin.ch/ch/d/ff/2001/29.pdf (besucht am 17. Januar 2005)
- Bundesamt für Statistik: Medienmitteilung vom 10. Mai 1996
- Bundesamt für Statistik: Medienmitteilung vom Februar 1999; erhältlich unter www.bfs.admin.ch/bfs/portal/de/index/news/medienmitteilungen.html (besucht am 24. März 2005)
- Bundesamt für Statistik: Strassenverkehrsunfälle in der Schweiz 2003, Neuchâtel 2004
- Bundesamt für Statistik: Strafurteilsstatistik 2004; erhältlich unter www.bfs.admin.ch/bfs/portal/de/index/themen/rechtspflege/kriminalitaet__strafvollzug.html (besucht am 17. Februar 2005)
- Bundesrat: Bericht zum Postulat Ruffy, Sterbehilfe, Ergänzung des Strafgesetzbuchs, vom Juli 2000; erhältlich unter www.ofj.admin.ch/themen/stgb-sterbehilfe/ber-ruffy-d.pdf (besucht am 3. Februar 2005)
- Eidgenössische Steuerverwaltung, Abteilung Steuerstatistik und Dokumentation: Direkte vs. Indirekte Steuern: Einige Bemerkungen zum Steuermix in der Schweiz, Bern 2003; erhältlich unter [www.estv.admin.ch/data/d/paket/ tax_mix.pdf](http://www.estv.admin.ch/data/d/paket/tax_mix.pdf) (besucht am 29. Januar 2005)
- Justizvollzug des Kantons Zürich: Jahresheft 2004; erhältlich unter [www.poeschwies.ch/content/ pdf/suche/index.php?id_pdf=395](http://www.poeschwies.ch/content/pdf/suche/index.php?id_pdf=395) (besucht am 10. Januar 2005)
- Kantonsgericht Graubünden: Jahresbericht 2003; erhältlich unter [www.kg-gr.ch/ -> Aktuelles/ Infos](http://www.kg-gr.ch/->Aktuelles/Infos) (besucht am 29. Januar 2005)
- National Research Council: The Evaluation of Forensic DNA Evidence, Washington DC 1996
- Obergericht des Kantons Aargau: Rechenschaftsbericht Justizbehörden 2001/2002; erhältlich unter www.ag.ch/obergericht/de/pub/angebote/rechenschaftsberichte.htm (besucht am 29. Januar 2005)
- Obergerichts des Kantons Bern: Verwaltungsbericht 2003; erhältlich unter [www.be.ch/OG/ -> Publikationen](http://www.be.ch/OG/->Publikationen) (besucht am 29. Januar 2005)

Allgemeiner Teil

Schwedische Akademie für Wissenschaften: Pressemitteilung vom 13. Oktober 1992; erhältlich unter <http://nobelprize.org/economics/laureates/1992/press.html> (besucht am 30. März 2005)

ALLGEMEINER TEIL

§ 1 Einleitung, Aufbau der Arbeit und Eingrenzung des Themas

I. Einleitung

- 1 Juristen und Soziologen¹ nehmen schon lange an, dass Richter nicht ausschliesslich rechtlichen Prinzipien folgen, wenn sie Fakten feststellen und Sachverhalte unter Normen subsumieren. Manche gehen davon aus, dass politische Ideologie das richterliche Urteil (mit-) bestimmt,² andere vermuten Rücksicht auf existierende Machtstrukturen,³ Grundüberzeugungen,⁴ das Geschlecht des Urteilenden⁵ oder die Förderung der eigenen Karriere.⁶
- 2 Die Forschung zu den so genannten *heuristics and biases* legt nahe, dass – unbeeinflusst von Ideologie und Eigennutz – die Natur menschlichen Denkens in Entscheidungssituationen unter Unsicherheit dazu führt, dass Richter in bestimmten Situationen systematische und vorhersehbare Fehler machen.⁷ AMOS TVERSKY und DANIEL KAHNEMAN zeigten in den frühen siebziger Jahren, dass Menschen bei Entscheidungen unter Unsicherheit so genannte Urteilsheuristiken – einfache Faustregeln, die komplexe Urteile vereinfachen – verwenden.⁸ Die Verwendung solcher Faustregeln kann häufig einen schwierigen Entscheid vereinfachen.⁹ In anderen Situationen führen sie jedoch zu systematischen Verzerrungen und Abweichungen von einer rationalen Entscheidung.¹⁰ Weil der Einfluss der *heuristics* und die daraus resultierenden *biases* von den Urteilenden meist nicht erkannt

¹ Frauen sind jeweils mitgemeint. Ein erster Entwurf dieser Arbeit verwendete abwechselungsweise das männliche und das weibliche Geschlecht, was jedoch bei Testlesern auf ein negatives Echo stiess.

² JEROME FRANK, *Courts on Trial: Myth and Reality in American Justice*, Princeton 1949.

³ DOROTHEE PETERS, *Richter im Dienste der Macht: Zur gesellschaftlichen Verteilung der Kriminalität*, Stuttgart 1973.

⁴ JEFFREY A. SEGAL/HAROLD J. SPAETH, *Supreme Court and the Attitudinal Model Revisited*, Cambridge 2002.

⁵ GREGORY C. SISK/MICHAEL HEISE/ANDREW P. MORRIS, *Charting the Influences on the Judicial Mind: An Empirical Study of Judicial Reasoning*, *New York University Law Review* 1998, 1377-1421, 1385 ff.

⁶ RICHARD A. POSNER, *What Do Justices Maximize? (The Same Thing As Everybody Else Does)*, *Supreme Court Economic Review* 1993, 1-28.

⁷ Für einen Überblick über den aktuellen Forschungsstand siehe THOMAS GILOVICH/DALE GRIFFIN/DANIEL KAHNEMAN (Hrsg.), *Heuristics and Biases: The Psychology of Intuitive Judgment*, Cambridge 2002.

⁸ AMOS TVERSKY/DANIEL KAHNEMAN, *Judgment under Uncertainty: Heuristics and Biases*, *Science* 1974, 1124-1131.

⁹ GERD GIGERENZER/JEAN CZERLINSKI/LAURA MARTIGNON, *How Good Are Fast and Frugal Heuristics?*, in: GILOVICH/GRIFFIN/KAHNEMAN (Hrsg.), 559-581.

¹⁰ Rational im Sinne der *expected utility* Theorie, formuliert von JOHN VON NEUMANN/OSKAR MORGENTHAU, *Theory of Games and Economic Behavior*, 2. Auflage, Princeton 1947.

werden, werden sie in Anlehnung an die bekannten optischen Täuschungen auch als „kognitive Täuschungen“ (*cognitive illusions*) bezeichnet.¹¹

- 3 Entscheidend ist, dass die erstmals von KAHNEMAN und TVERSKY systematisch untersuchten Urteilsheuristiken nicht zu zufälligen, sondern zu systematischen Verzerrungen des Urteilens führen. Dass der Mensch irrt, ist nichts Neues. Irrt er aber systematisch, so kann er versuchen, dem Irrtum entgegenzuwirken. Sein Irrtum kann aber auch von Dritten ausgenutzt werden.¹²
- 4 Die Erkenntnisse von KAHNEMAN und TVERSKY hatten und haben grossen Einfluss auf die Wirtschaftswissenschaft.¹³ Mit der *Prospect Theory* formulierten sie ein Modell menschlichen Entscheidungsverhaltens, welches tatsächliche Entscheidungen unter Unsicherheit besser abbildet als traditionelle Modelle.¹⁴ 2002 erhielt DANIEL KAHNEMAN für seine Forschung den Nobelpreis für Wirtschaftswissenschaften.¹⁵
- 5 Seit etwa Mitte der neunziger Jahre des letzten Jahrhunderts ist in den USA auch die Rechtswissenschaft auf das Forschungsprogramm von KAHNEMAN und TVERSKY aufmerksam geworden und versucht, das neue Menschenbild gewinnbringend zur Analyse des Verhaltens der Akteure im Rechtssystem anzuwenden. Der Ansatz wird gemeinhin als „*Behavioral Law and Economics*“ (BLE) bezeichnet, was bereits andeutet, dass er sich (zumindest ursprünglich) als Weiterentwicklung der traditionellen Ökonomischen Analyse des Rechts (ÖAR) versteht.¹⁶ „*Behavioral*“ sollte dabei nicht mit „behavioristisch“ übersetzt werden. „Behavioristisch“ bezeichnet im Deutschen eine psychologische Schule, die von JOHN B. WATSON (1878 -1958) und EDWARD LEE THORNDIKE (1874-1949) begründet wurde und nur das äusserlich sicht- und messbare Verhalten als legitimen Gegenstandsreich der Psychologie sieht. Der *heuristics and biases* Ansatz steht hingegen in der Tradition der „kognitiven Wende“ der 60-er Jahre, die unter bewusster Abgrenzung vom klassischen Behaviorismus (erneut) innere Prozesse (Denken, Fühlen) als legitimen Gegenstand wissenschaftlicher psychologischer Forschung ansieht.¹⁷ Wenn man „*Behavioral Law and Economics*“ ins Deutsche übersetzen will, so am Besten als „Verhaltenswissenschaftliche Ökonomie und Recht“.

¹¹ WOLFGANG HELL, Kognitive und optische Täuschungen, in: WOLFGANG HELL/KLAUS FIEDLER/GERD GIGERENZER (Hrsg.), Kognitive Täuschungen: Fehl-Leistungen und Mechanismen des Urteilens, Denkens und Erinnerns, Heidelberg 1993, 317-324.

¹² Nichts anderes versuchen Fondsmanager, die das ihnen anvertraute Vermögen nach Erkenntnissen der *behavioral finance* anlegen; der Bekannteste unter ihnen dürfte der Chicagoer Professor RICHARD H. THALER sein (Fuller & Thaler Asset Management, Inc.).

¹³ DAVID LAIBSON/RICHARD ZECKHAUSER, Amos Tversky and the Ascent of Behavioral Economics, *Journal of Risk and Uncertainty* 1998, 7-47; ERNST FEHR/GERHARD SCHWARZ, Psychologische Grundlagen der Ökonomie, Zürich 2002, 10.

¹⁴ DANIEL KAHNEMAN/AMOS TVERSKY, Prospect theory: An analysis of decision under risk, *Econometrica* 1979, 263-291.

¹⁵ www.nobel.se/economics/laureates/2002/index.html (besucht am 2. Mai 2004). AMOS TVERSKY verstarb 1996; der Nobelpreis wird nicht posthum verliehen.

¹⁶ CHRISTINE JOLLS/CASS R. SUNSTEIN/RICHARD H. THALER, A Behavioral Approach to Law and Economics, *Stanford Law Review* 1998, 1471-1550; Nachdruck in: SUNSTEIN (Hrsg.), 13-58.

¹⁷ JEFFREY J. RACHLINSKI, The „New“ Law and Psychology: A Reply to Critics, Sceptics and Cautious Supporters, *Cornell Law Review* 2000, 739-766, 740.

- 6 In Europa hat die neue Forschungsrichtung bisher keine grossen Wellen geworfen, aber mit der üblichen Verspätung zeichnet sich jetzt auch hier zunehmendes Interesse ab. Im Mai 2004 hat die 94. Dahlem Konferenz der Freien Universität Berlin zum Thema „*Heuristics and the Law*“ stattgefunden, an der die führenden Vertreter der Richtung teilgenommen haben.¹⁸ Ich bin selbst erstmals 2001 anlässlich eines Seminars von Prof. JAMES KRIER und RONAN AVRAHAM an der *University of Michigan Law School* mit dem Forschungsansatz der Verhaltenswissenschaftlichen Ökonomischen Analyse des Rechts in Kontakt gekommen. Fasziniert von diesem Ansatz, mit dem ich während meiner Ausbildung in der Schweiz nie konfrontiert worden war, schrieb ich meine *Master-Thesis* über den so genannten „Besitztumseffekt“ (mehr dazu hinten, S. 58 ff.), wofür ich ein Experiment mit 40 Studierenden durchführte.¹⁹ Was ich dabei über experimentelle Forschung und Statistik lernte, half mir, die vorliegende Arbeit anzupacken.
- 7 Inspiriert von einer Studie von GUTHRIE, RACHLINSKI und WISTRICH führte ich im Herbst 2003 eine Umfrage unter schweizerischen Richterinnen und Richtern durch.²⁰ Ziel war, mittels eines Aufsatzes Schweizer Juristen auf das Phänomen der kognitiven Täuschungen aufmerksam zu machen. Nach der Auswertung der Umfrage und der Redaktion des Artikels war einerseits viel Zeit verstrichen, andererseits hatte der Artikel einen Umfang angenommen, der eine Publikation in einer Zeitschrift unwahrscheinlich machte. Ich beschloss deshalb, die Arbeit zu einer Dissertation auszubauen und führte im Herbst 2004 eine weitere Umfrage mit (anderen) Schweizer Richtern durch. Die Ergebnisse der beiden Umfragen bilden das Herzstück des zweiten Teils dieser Arbeit.

II. Aufbau der Arbeit und Eingrenzung des Themas

- 8 Im ersten Teil werden nach der Einleitung sehr kurz Geschichte und Grundlagen der Ökonomischen Analyse des Rechts dargestellt. Dies ist notwendig, weil *Behavioral Law and Economics* als Kritik auf die (in den USA) dominierende traditionelle Ökonomische Analyse des Rechts entstanden und ohne diesen Hintergrund kaum verständlich ist. Da Vertreter der BLE seit neustem versuchen, sich – von der ÖAR abgrenzend – in die Tradition des amerikanischen Rechtsrealismus zu stellen, beginnt der geschichtliche Rückblick mit einer Skizze des *American Legal Realism*. Anschliessend wird nach einem Überblick über die wichtigsten Arbeiten der jungen Disziplin auf die Kritik am *heuristics and biases* Ansatz in der Psychologie und die Kritik an der BLE in der Jurisprudenz eingegangen.
- 9 Im zweiten Teil werden die Methode der Studie und Gedanken zur externen Validität erläutert und die Resultate der beiden Umfragen dargestellt, die als Aufhänger für eine „*tour d’horizon*“ der einzelnen Täuschungen dienen. Dabei wird einerseits der psychologi-

¹⁸ Der von GERD GIGERENZER und CHRISTOPH ENGEL herausgegebene Tagungsband zur 94. Dahlem Konferenz konnte leider nicht mehr verarbeitet werden, da er zur Zeit der Abgabe dieser Arbeit (Anfang Mai 2005) noch nicht vorlag.

¹⁹ MARK SCHWEIZER, *The Offer/Asking Price Gap and Term of Possession – an Experiment*, unpublizierte Master-These Universität Michigan, 2002 (erhältlich unter www.decisions.ch/publikationen.html; besucht am 1. Mai 2005).

²⁰ CHRIS GUTHRIE/JEFFREY J. RACHLINSKI/ANDREW J. WISTRICH, *Inside the Judicial Mind*, *Cornell Law Review* 2001, 777-830.

schen Forschung zu den einzelnen kognitiven Täuschungen relativ viel Raum eingeräumt, da die entsprechende Literatur für Juristen in der Schweiz nur schwer zugänglich ist und kaum bekannt sein dürfte. Andererseits werden die Auswirkungen der kognitiven Täuschungen auf juristische Entscheidungen erörtert, wobei hier das Thema auf Entscheidungen von Richtern und Parteien in Gerichtsverfahren beschränkt ist. Selbstverständlich unterliegen auch der Gesetzgeber und die Rechtsunterworfenen kognitiven Täuschungen, und zahlreiche Artikel der BLE beschäftigen sich mit den Auswirkungen kognitiver Täuschungen auf diese Personengruppen. Sie sind aber nicht Gegenstand dieser Arbeit.

- 10 Aus psychologischer Sicht beschränkt sich die Arbeit in erster Linie auf die Darstellung des Einflusses der von KAHNEMAN und TVERSKY erforschten Urteilsheuristiken auf Richter und Parteien in Gerichtsverfahren. Daneben werden einzelne weitere kognitive Phänomene (wie Kompromiss- und Kontrasteffekt, S. 243 ff.) vorgestellt. Selbstverständlich können emotionale und motivationale Faktoren das richterliche Urteil beeinflussen – nicht umsonst soll der Richter sein Urteil *sine ira et studio* fällen. Solche Faktoren bilden aber auch nicht Gegenstand dieser Arbeit. Ausgeklammert bleibt schliesslich die Forschung zum Fairness-Empfinden und (echten) Altruismus der Menschen, die ebenfalls im Widerspruch zu den Annahmen der klassischen ÖAR stehen.²¹

²¹ Siehe JOSEPH HENRICH/ROBERT BOYD/SAMUEL BOWLES/COLIN CAMERER/ERNST FEHR/HERBERT GINTIS, *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*, New York 2004.

§ 2 American Legal Realism

- 11 Die amerikanischen Rechtsrealisten (*legal realists*) wandten sich in den zwanziger und dreissiger Jahren des 20. Jahrhunderts gegen den damals vorherrschenden „Formalismus“ und die „mechanische Jurisprudenz“ der klassischen Juristen. Die Rechtsrealisten bezweifelten, dass Streitfälle aufgrund von Rechtsregeln entschieden wurden und bestritten, dass die „Lösung“ von rechtlichen Problemen aus abstrakten Konzepten abgeleitet werden konnte. Die Rechtsrealisten selber betrachten OLIVER WENDELL HOLMES, JR. (1841-1935) als ihren intellektuellen Vorfahren, der Rechtsrealismus als eigenständige intellektuelle Strömung aber entstand in den zwanziger Jahren des letzten Jahrhunderts an den Rechtsfakultäten von Columbia und Yale. KARL N. LLEWELLYN, WALTER WHEELER COOK, UNDERHILL MOORE und HERMAN OLIPHANT waren ihre wichtigsten Vertreter; JEROME FRANK in mancher Hinsicht ihr extremster. Als „amerikanischer“ Rechtsrealismus wird die Strömung in Abgrenzung zum ungefähr gleichzeitig blühenden skandinavischen Rechtsrealismus bezeichnet, auf den ich hier nicht weiter eingehe.²²
- 12 Die Rechtsrealisten argumentieren, dass das Recht in zweifacher Hinsicht unbestimmt (*indeterminate*) ist.²³ Erstens ist das Recht *normativ* unbestimmt in dem Sinne, dass die verfügbaren rechtlichen Argumente und Konzepte ein bestimmtes Urteil nicht *rechtfertigen* können; zumindest nicht in den schwierigen Fällen (*hard cases*), wie sie in der Regel vor Berufungsgerichten verhandelt werden.²⁴ Zweitens ist das Recht *deskriptiv* unbestimmt in dem Sinne, dass die rechtlichen Regeln nicht genügen, um zu erklären, warum ein Gericht einen Fall so und nicht anders entschieden hat. LLEWELLYN wies darauf hin, dass Präjudizien eng oder weit interpretiert werden können und es kein logisches Argument gibt, der einen oder anderen Interpretation den Vorzug zu geben.²⁵ Bei einer engen Interpretation wird die bindende Wirkung des Präjudizes auf die spezifischen Fakten beschränkt, die dem Fall zu Grunde lagen, während bei der weiten Interpretation versucht wird, ein allgemein gültiges Prinzip abzuleiten, das auf alle vergleichbaren Fälle anwendbar ist. Präjudizien als Rechtsquellen können daher ein Urteil allein nicht erklären, da verschiedene Regeln aus ihnen abgeleitet werden können.
- 13 Da sich die Realisten einig waren, dass sich aus Rechtsregeln alleine ein Gerichtsurteil in einem schwierigen Fall nicht voraussagen lässt, musste es etwas anderes sein, das den Ausgang eines Verfahrens bestimmte. Für BRIAN LEITER ist Kern (*core claim*) der realistischen Sicht, dass Richter in erster Linie auf die dem Fall zugrunde liegenden *Fakten*

²² Für einen Überblick über Gemeinsamkeiten und Unterschiede der beiden Realismen siehe GREGORY S. ALEXANDER, Comparing the Two Legal Realisms – American and Scandinavian, American Journal of Comparative Law 2002, 131-174.

²³ Die Darstellung folgt dem Aufbau von BRIAN LEITER, American Legal Realism, in: WILLIAM EDMUNDSON/ MARTIN GOLDING (Hrsg.), The Blackwell Guide to Philosophy of Law and Legal Theory, Oxford 2004, 50-66.

²⁴ KARL LLEWELLYN, Some Realism about Realism – Responding to Dean Pound, Harvard Law Review 1931, 1222-1264, 1239 (die Unbestimmtheit des Rechts einschränkend auf Fälle, die von vernünftigen Parteien durchprozessiert werden: in *diesen* Fällen gibt es in der Regel mindestens zwei widersprüchliche *vertretbare* Interpretationen des anwendbaren Rechts).

²⁵ KARL N. LLEWELLYN, The Bramble Bush, New York 1930, 74 ff.

reagieren, wenn sie einen Fall entscheiden, und nicht auf rechtliche Normen und Konzepte.²⁶ HERMAN OLIPHANT schrieb, dass Gerichte auf den Stimulus der Fakten des spezifisch zu entscheidenden Falles reagieren und nicht auf den Stimulus der allzu generellen und abgenutzten abstrakten Regeln, die in Urteilen und Kommentaren postuliert werden.²⁷ JEROME FRANK zitiert einen angesehenen Richter, der ihm gestanden habe, er sehe aufgrund der Fakten, wo die Gerechtigkeit liege. Erst dann setze er sich hin und suche nach Rechtsquellen, und in fast allen Fällen finde er Prinzipien, die seiner Sicht des Falles entsprächen.²⁸ LLEWELLYNS Rat an Anwälte, das Gericht aufgrund des Sachverhaltes zu überzeugen, dass die von ihnen vertretene Partei zu gewinnen verdient und dem Gericht bloss eine „technische Brücke“ zur Verfügung zu stellen, die das Urteil rechtfertigt, setzt die gleiche Sicht voraus.²⁹

- 14 Die Aussage allein, dass das Gericht aufgrund der Fakten entscheidet, ist natürlich banal. Die Aussage der Realisten ist stärker: sie behaupten, dass Gerichte (auch) aufgrund von Fakten entscheiden, die gemäss den anwendbaren Normen *irrelevant* sind.³⁰ Ein Ziel der Rechtsrealisten war daher, die ihrer Ansicht nach zu abstrakten rechtlichen Prinzipien neu zu formulieren, so dass sie die Fakten, die den Urteilen *tatsächlich* zu Grunde liegen, berücksichtigen.³¹
- 15 Die Behauptung, dass Richter Fälle aufgrund der Fakten und nicht der Normen entschieden, wirft natürlich die Frage auf, welche Fakten entscheidend sind. Der „soziologische“ Flügel der Rechtsrealisten – OLIPHANT, MOORE und LLEWELLYN – dachte, dass juristische Entscheidungen voraussehbar sind: vorauszusagen nicht (nur) durch das Studium der Rechtsquellen, sondern durch das Studium der sozialen Kräfte, die auf Richter wirken. Der von LEITER als „idiosynkratisch“ bezeichnete Flügel hingegen behauptete, dass Persönlichkeit und Psychologie des Richters determinieren, wie er auf die Fakten reagiert.³² Für FRANK, den prominentesten Vertreter des idiosynkratischen Flügels, bestimmen unbewusste Begehren die Entscheidung des Richters; da diese für einen aussenstehenden Beobachter kaum zu erkennen sind, sind richterliche Entscheidungen unmöglich zu prognostizieren. Kritiker der Realisten machten sich über diese Sicht lustig, indem sie behaupteten, gemäss den Realisten bestimme „was der Richter zum Frühstück hatte“ den Ausgang von Gerichtsverfahren; eine Aussage, die FRANK so nie gemacht hat.³³
- 16 Vertreter des soziologischen Flügels glauben, dass das Studium des sozialen Kontexts von juristischen Entscheidungen, der Einstellung von Richtern und ihrer Berufsauffassung die Voraussehbarkeit von Gerichtsurteilen verbessern kann.³⁴ Rechtssoziologie ist daher für

²⁶ LEITER, FN 23, 52.

²⁷ HERMAN OLIPHANT, A Return to Stare Decisis, American Bar Association Journal 1928, 71-76, 107, 159-162, 75. Der Begriff „Stimulus“ lehnt sich bewusst an die Terminologie des behavioristischen Psychologen WATSON an.

²⁸ JEROME FRANK, Law and the Modern Mind, New York 1930, 104.

²⁹ KARL LLEWELLYN, FN 25, 76.

³⁰ LEITER, FN 23, 53.

³¹ OLIPHANT, FN 27, 159 f.

³² LEITER, FN 23, 54.

³³ LEITER, FN 23, 54.

³⁴ JOSEPH WILLIAM SINGER, Legal Realism Now, California Law Review 1988, 467-544, 470 f.

LLEWELLYN in erster Linie eine empirische, deskriptiv-analytische Wissenschaft. Wenn das Recht sich im Verhalten des Rechtsstabes äussert – in Anlehnung an den viel zitierten Ausspruch von HOLMES „*the prophecies of what the courts will do in fact and nothing more pretentious are what I mean by the law*“³⁵ – dann muss eine Wirklichkeitswissenschaft vom Recht Verhaltensforschung sein.³⁶ Die empirischen Studien der Rechtsrealisten – die als erste Juristen selber empirisch forschten – werden aber wegen ihrer Theorie- und Konzeptlosigkeit als wenig erfolgreich angesehen.³⁷

- 17 Da Rechtsregeln gerichtliche Entscheidungen nach Ansicht der Rechtsrealisten nicht nur unvollständig beschreiben, sondern auch nicht alleine rechtfertigen können, müssen Urteile durch (nach klassischem Verständnis) ausserrechtliche Argumente begründet werden. Die Rechtsrealisten appellierten daher an die Richter, die Folge- und Zweckmässigkeitsüberlegungen (*policy considerations*), die ihren Urteilen implizit ohnehin zu Grunde lägen, explizit offen zu legen und damit zur Diskussion zu stellen.³⁸ LEITER nennt die Rechtsrealisten, die diese Forderung aufstellen (v.a. HOLMES, FELIX COHEN und FRANK als Richter) „*Proto-Posnerian*“, weil RICHARD POSNER, der wohl bekannteste Vertreter der Ökonomischen Analyse des Rechts, fordert, dass Gerichtsurteile explizit mit Effizienzüberlegungen (und damit Folge- und Zweckmässigkeitsüberlegungen) begründet werden sollen.³⁹ Folgeüberlegungen setzen wiederum empirische sozialwissenschaftliche Studien, oder zumindest eine Theorie sozialen Verhaltens (wie die Ökonomik), voraus, wenn sie überzeugend sein sollen. Damit spielen die Sozialwissenschaften auch eine Rolle bei der Frage, wie das Recht sein soll, und nicht nur bei der Frage, warum es so ist, wie es ist.
- 18 Dem deutschsprachigen Juristen fallen die Parallelen des amerikanischen Rechtsrealismus zur (zeitlich früheren) deutschen Freirechtsschule (EUGEN EHRLICH, ERNST FUCHS, HERMAN KANTOROWICZ) auf. EHRLICH, der von den Freirechtlern als geistiges Haupt anerkannt wurde,⁴⁰ hatte in „*Freie Rechtsfindung und freie Rechtswissenschaft*“ bereits 1903 die technische juristische Methode kritisiert. Er verlangte, dass sich die Rechtswissenschaft von dogmatischen Klimmzügen („unfruchtbarer Scharfsinn“) weg- und dem Studium des Rechts in seinem sozialen Kontext zuwendet.⁴¹
- 19 Die Parallelen zwischen *Legal Realism* und Freirechtsschule sind kein Zufall. ROSCOE POUND, der den Begriff der „soziologischen Jurisprudenz“ prägte, scheint die meisten

³⁵ OLIVER WENDELL HOLMES, JR., *The Path of the Law*, Harvard Law Review 1897, 457- 469, 460; erhältlich unter www.gutenberg.org/etext/2373 (besucht am 20. April 2005).

³⁶ MANFRED REHBINDER, Karl N. Llewellyn als Rechtssoziologe, *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 1966, 532-556, 546.

³⁷ JOHN H. SCHLEGEL, *American Legal Realism and Empirical Social Science*, Durham 1995, 233 f.; M.D.A. FREEMAN, *Lloyd's Introduction to Jurisprudence*, 7. Aufl. London 2001, 803.

³⁸ So bereits HOLMES, FN 35, 467.

³⁹ LEITER, FN 23, 58.

⁴⁰ MANFRED REHBINDER, Neues über Leben und Werk von Eugen Ehrlich, in: FRIEDRICH KAULBACH/WERNER KRAWIETZ (Hrsg.), *Festschrift für Helmut Schelsky zum 65. Geburtstag*, Berlin 1978, 403-418, 408.

⁴¹ EUGEN EHRLICH, *Freie Rechtsfindung und freie Rechtswissenschaft*, Leipzig 1903; Nachdruck in: EUGEN EHRLICH, *Recht und Leben: Gesammelte Schriften zur Rechtstatsachenforschung und zur Freirechtslehre*, ausgewählt und eingeleitet von MANFRED REHBINDER, Berlin 1967, 170-202.

Werke der Freirechtler gelesen zu haben und hat ihre Ideen in seinen frühen Werken für die USA rezipiert.⁴² Er meinte: „we have a great deal of freie Rechtsfindung in America, while disclaiming it in theory“.⁴³ 1917 erschien „*The Science of Legal Method*“, ein Sammelband mit englischen Übersetzungen der wichtigsten Texte der Freirechtler, darunter FRANÇOIS GENYS „*Méthode d'interprétation et sources en droit privé positif*“ und EHRLICHs „Freie Rechtsfindung“.⁴⁴ BENJAMIN CARDOZOS einflussreiches Werk „*The Nature of Judicial Process*“, publiziert 1921, zitiert GENYS „*Méthode*“ elf Mal, EHRLICH 16 Mal und KANTOROWICZ' „Kampf um die Rechtswissenschaft“ sieben Mal.⁴⁵ LLEWELLYN schliesslich gab zwischen 1928 und 1932 Gastvorlesungen in Leipzig, beabsichtigte, EHRLICHs „Grundlegung der Soziologie des Rechts“ ins Englische zu übersetzen und war ein Freund von KANTOROWICZ, mit dem er einen regen Briefverkehr führte.⁴⁶

- 20 Trotzdem wird der Rechtsrealismus von Amerikanern als einziger eigenständiger Beitrag der USA zu einer generellen Rechtstheorie bezeichnet.⁴⁷ Die Rechtsrealisten beriefen sich nie explizit auf deutsche Quellen (auch nicht auf deren englische Übersetzungen), so dass dieser Eindruck leicht entstehen konnte. Der Grund dafür dürfte darin zu suchen sein, dass Deutschland und Deutsche in den zwanziger und dreissiger Jahren des letzten Jahrhunderts in den USA, um es gelinde zu sagen, nicht besonders populär waren.⁴⁸
- 21 Nach dem zweiten Weltkrieg galt der Rechtsrealismus in den USA als „tot“. Dafür waren verschiedene Gründe verantwortlich. Einerseits gelang es den Realisten nicht, ein fruchtbares empirisches Forschungsprogramm auf die Beine zu stellen;⁴⁹ andererseits bezog die Strömung (wie auch die Freirechtler) einen grossen Teil ihrer Kraft aus der Kritik an der herrschenden Jurisprudenz. Es war relativ einfach, dogmatische Argumente der klassischen *case law* Methode oder der Begriffsjurisprudenz als hohl zu entlarven, aber schwierig, etwas Neues an deren Stelle zu errichten. In dieser Hinsicht versagten die Realisten in den Augen mancher Rechtshistoriker; HERGET/WALLACE meinen: „*Much was torn down, little was built*“.⁵⁰ Den Todesstoss versetzte dem *Legal Realism* der englische Rechtsphilosoph HLA HART in seinem 1961 publizierten, einflussreichen Buch „*The Concept of Law*“, mit dem er den Positivismus wieder belebte. Ein Kapitel war den „Regel-Skeptikern“ gewidmet, womit HART die Rechtsrealisten meinte. Obwohl die Kritik von HART an den Rechtsrealisten nicht besonders überzeugend ist und auf einer irreführenden Beschreibung der

⁴² JAMES E. HERGET/STEPHEN WALLACE, *The German Free Law Movement as the Source of American Legal Realism*, *Vanderbilt Law Review* 1987, 399-455, 422.

⁴³ ROSCOE POUND, *Enforcement of Law*, *Green Bag* 1908, 401-415, 406 (zitiert nach HERGET/WALLACE, FN 42, 422).

⁴⁴ HERGET/WALLACE, FN 42, 428.

⁴⁵ HERGET/WALLACE, FN 42, 428.

⁴⁶ ULRICH DROBNIG, Llewellyn and Germany, in: ULRICH DROBNIG/MANFRED REHBINDER (Hrsg.), *Rechtsrealismus, multikulturelle Gesellschaft und Handelsrecht: Karl N. Llewellyn und seine Bedeutung heute*, Berlin 1994, 17-43, 18, 29 f. und 35 f.

⁴⁷ ROBERT S. SUMMERS, *Instrumentalism and American Legal Theory*, Ithaca 1982, 19 (zitiert nach HERGET/WALLACE, FN 42, 400).

⁴⁸ HERGET/WALLACE, FN 42, 432.

⁴⁹ SCHLEGEL, FN 37, 213 f.

⁵⁰ HERGET/WALLACE, FN 42, 435.

realistischen Position beruht,⁵¹ führte sie dazu, dass sich die Rechtsphilosophen vom Rechtsrealismus abwandten.⁵²

- 22 Die Rechtsrealisten hinterliessen aber ein einflussreiches Erbe, nämlich ihr instrumentelles Rechtsverständnis. Um es mit JOSEPH W. SINGER zu sagen:⁵³

Es ist nicht länger möglich, der Forderung, dass Vermieter für alle Schäden der Mieter kausal haften, mit dem Argument entgegenzutreten: „es liegt in der Natur der Miete, dass sie Besitz überträgt und daher der Mieter, als unmittelbarer Besitzer der Mietsache, für sich selber verantwortlich ist“. Obwohl das 1890 übliches juristische Denken war, überzeugt es heute schlicht nicht mehr als rechtliches Argument. Der rechtliche Diskurs hat sich wegbewegt von der Deduktion von Folgen aus abstrakten Konzepten hin zum Versuch, das Recht durch Zweckmässigkeitsüberlegungen (*policy*), Moral und institutionellen Interessen zu rechtfertigen. Dieser revolutionäre Wandel im juristischen Diskurs stellt eine monumentale Leistung dar.

- 23 Die Vertreter der Ökonomischen Analyse des Rechts sehen die Aufgabe des Richters ebenfalls nicht darin, Fälle durch die logische Ableitung von Folgen aus abstrakten Konzepten zu entscheiden, sondern gemäss Regeln, die ein soziales Ziel – im Falle der Ökonomen „Effizienz“ – am besten verwirklichen. Vertreter des Effizienzzieles im Rechts glauben wie die Rechtsrealisten, dass eine juristische Doktrin an ihren sozialen Folgen gemessen werden muss. Ob eine Rechtsregel effizient ist, lässt sich nur sozialwissenschaftlich feststellen. Auch halten Vertreter der ÖAR die formalistische Herleitung von Normen für falsch, da sie die tatsächlichen Entscheidungsgründe vernebelt und eine offene Diskussion über die politischen Interessen, die jeder Entscheidung zu Grunde liegen, verhindert. In diesem Sinne kann die Ökonomische Analyse des Rechts als „Anwendung des Rechtsrealismus“ bezeichnet werden.⁵⁴ Andererseits ist die Ökonomische Analyse des Rechts formalistisch, weil sie auf einem formalen Menschenbild beruht, das empirischer Überprüfung nicht standhält, und die Gleichsetzung von „Nutzen“ im utilitaristischen Sinn mit der Bereitschaft zu zahlen ein formaler, wenig überzeugender Kunstgriff ist.⁵⁵

⁵¹ LEITER, FN 23, 61 ff.

⁵² LEITER, FN 23, 65.

⁵³ SINGER, FN 34, 475.

⁵⁴ SINGER, FN 34, 515.

⁵⁵ SINGER, FN 34, 522 ff.

§ 3 Ökonomische Analyse des Rechts

- 24 Die Ökonomik – d.h. die wirtschaftswissenschaftliche *Methode*, im Gegensatz zur Ökonomie, ihrem traditionellen Gegenstand – beschäftigt sich mit zwei unterschiedlichen Fragestellungen. Die positive Ökonomik erklärt und prognostiziert wirtschaftliche Vorgänge. Die normative Ökonomik hingegen bewertet wirtschaftliche Zustände oder Veränderungen auf der Grundlage eines vorher definierten Kriteriums.⁵⁶ Die Ökonomische Analyse des Rechts als Teilgebiet der Ökonomik beschäftigt sich daher sowohl mit positiven wie normativen Fragestellungen, die nicht vermischt werden sollten.⁵⁷
- 25 Zusammenfassend ergeben sich folgende Thesen der ÖAR:⁵⁸
- Positiv: Menschen reagieren auf rechtlichen Zwang wie *hominī oeconomici*, also rational und nutzenmaximierend. Rechtliche Sanktionen wirken wie Preise, die bestimmte Verhaltensweisen gegenüber anderen verteuern. Die Voraussage der sozialen Steuerungswirkung des Rechts erfolgt auf der Grundlage des ökonomischen Verhaltensmodells.
- Normativ: Eine Rechtsnorm ist gut, wenn sie den gesellschaftlichen Wohlstand erhöht, schlecht, wenn sie ihn verringert. Eine Bewertung der durch Rechtsnormen ausgelösten Folgen erfolgt auf Grundlage der Wohlfahrtsökonomik, insbesondere des Kaldor/Hicks-Kriteriums. Das Kaldor/Hicks-Kriterium, auch als potentielle Pareto-Verbesserung bezeichnet, besagt, dass ein Zustand X der Welt besser ist als ein Zustand Y, wenn die Vorteile der Gewinner bei Zustand X so gross sind, dass sie die Verlierer kompensieren *könnten* – ungeachtet dessen, ob die Kompensation tatsächlich stattfindet.⁵⁹ Um feststellen zu können, ob das Kaldor/Hicks-Kriterium erfüllt ist, muss eine Kosten-/Nutzenanalyse durchgeführt werden. Trotz theoretischer und praktischer Bedenken gegenüber dem Kaldor/Hicks-Kriterium ist es unverzichtbar für die angewandte Wohlfahrtsökonomik.⁶⁰

I. Ursprünge der Ökonomischen Analyse des Rechts

- 26 Obwohl es frühe Vorläufer gegeben hat,⁶¹ kann man den eigentlichen Ursprung der Ökonomischen Analyse des Rechts in den USA auf die sechziger Jahre des letzten Jahrhunderts datieren.⁶² Die Arbeiten von RONALD COASE, GUIDO CALABRESI, GARY S. BECKER und RICHARD POSNER gelten als die Grundsteine der ÖAR.

⁵⁶ HORST EIDENMÜLLER, Effizienz als Rechtsprinzip, Tübingen 1995, 21.

⁵⁷ GUNNAR JANSON, Ökonomische Theorie im Recht, Berlin 2004, 22.

⁵⁸ Nach EIDENMÜLLER, FN 56, 4.

⁵⁹ EIDENMÜLLER, FN 56, 51.

⁶⁰ ROBERT COOTER/THOMAS ULEN, Law and Economics, 4. Aufl., Reading 2003, 49.

⁶¹ Zu diesen JANSON, FN 57, 20.

⁶² GEBHARD KIRCHGÄSSNER, Homo Oeconomicus, 2. Aufl., Tübingen 2000, 134.

- 27 RONALD COASE führte 1961 mit seinem Aufsatz „*The Problem of Social Cost*“ die Transaktionskostenanalyse in die Rechtswissenschaft ein.⁶³ Das so genannte Coase-Theorem besagt, dass es keine Rolle spielt, wem das (Eigentums-)recht an einem Gut zugeordnet wird, wenn die Transaktionskosten gering sind. Das Gut wird letztendlich, über eine Reihe von freiwilligen Transaktionen, bei demjenigen landen, der ihm den höchsten Wert beimisst. Die Marktteilnehmer können somit Probleme, die durch externe Effekte (wie der Lärm einer Schneiderei oder der Funkenflug einer Eisenbahn in den Beispielen von COASE) auftreten, selber lösen, wenn sie über die Allokation der Ressourcen verhandeln und diese ohne Kosten transferieren können. Die Gültigkeit des Coase-Theorems wird durch den bereits erwähnten Besitztumseffekt in Frage gestellt: wenn die Präferenz für ein Gut davon abhängt, ob man es besitzt, spielt die ursprüngliche Zuteilung der Güter eben doch eine Rolle, und ein Gut wird auch in Abwesenheit von Transaktionskosten nicht notwendigerweise bei dem enden, der am meisten zu zahlen bereit ist. Der Besitztumseffekt wirft auch wohlfahrtsökonomische Probleme auf. Wer schätzt ein Gut mehr: derjenige, der mehr dafür zu zahlen bereit ist oder derjenige, der mehr für die Aufgabe des Gutes verlangt? M. a. W. ist die *willingness to pay* (WTP) oder *willingness to accept* (WTA) entscheidend für die Analyse?⁶⁴
- 28 GUIDO CALABRESI analysierte 1961 in seinem wichtigen Aufsatz „*Some Thoughts On Risk Distribution And The Law Of Torts*“ das ausserververtragliche Haftungsrecht.⁶⁵ CALABRESI'S Artikel ist von weniger grosser theoretischer Wichtigkeit als COASE'S „*Problem of Social Cost*“, aber er war der erste Versuch, ein ganzes Rechtsgebiet durch ökonomische Prinzipien zu erklären. CALABRESI wandte Einsichten aus der Ökonomik über Risiko und Versicherung auf das ausserververtragliche Haftungsrecht an, wobei er das Haftungsrecht zumindest teilweise als Versicherungssystem rekonzipierte. CALABRESI konnte demonstrieren, dass einige wenige ökonomische Prinzipien die Struktur des Haftungsrechts besser erklären können als zahlreiche überkommene rechtliche und moralische Regeln und Gebote. Diese Analyse des Haftungsrechts setzt unter anderem voraus, dass die Rechtsunterworfenen, der Gesetzgeber und die Rechtsprechenden Wahrscheinlichkeiten richtig wahrnehmen. Empirische Forschung, auf die sich Vertreter der *Behavioral Law and Economics* stützen, hat aber gezeigt, dass dies nicht der Fall ist und geringe Wahrscheinlichkeiten systematisch über-, mittlere bis hohe Wahrscheinlichkeiten aber unterschätzt werden (hinten, S. 60 ff.).
- 29 Es ist das Verdienst von GARY S. BECKER, als erster das ökonomische Modell auf Rechtsgebiete angewandt zu haben, die weit vom klassischen Gegenstand der Ökonomie entfernt sind. In „*Crime and Punishment: An Economic Approach*“ erklärte er 1968 die Wirkung von Strafnormen auf das Verhalten der Rechtsunterworfenen durch das ökonomische Modell.⁶⁶ Die Entscheidung, eine Straftat zu begehen, hängt nach diesem Modell von

⁶³ RONALD COASE, *The Problem of Social Cost*, Journal of Law and Economics 1960, 1-44.

⁶⁴ RUSSELL KOROBKIN, *Policymaking and the Offer/Asking Price Gap: Toward a Theory of Efficient Entitlement Allocation*, Stanford Law Review 1994, 663-708.

⁶⁵ GUIDO CALABRESI, *Some Thoughts On Risk Distribution And The Law Of Torts*, Yale Law Journal 1961, 499-533.

⁶⁶ GARY S. BECKER, *Crime and Punishment: An Economic Approach*, Journal of Political Economy 1968, 169-217.

der Abwägung der damit verbundenen Kosten und Nutzen ab. Die Kosten werden bestimmt durch die zu erwartende Strafe, abdiskontiert (d.h. multipliziert) mit dem Risiko, erwischt zu werden. Die Kosten einer Straftat können daher entweder durch härtere Strafen oder durch die Erhöhung des Risikos, erwischt zu werden, gesteigert werden. Vertreter der ÖAR haben dieses Modell benutzt, um für die Effizienz härterer Strafen einzutreten. Wenn, so POSNER, die Wahl besteht, den Täter mit einer Wahrscheinlichkeit von 0,1 zu erwischen und für zehn Jahre einzusperren oder ihn mit einer Wahrscheinlichkeit von 0,2 zu erwischen und für fünf Jahre einzusperren, dann sind bei der ersten Alternative halb so viele Personen doppelt so lange im Gefängnis wie bei der zweiten Alternative und die Kosten des Strafvollzugs daher dieselben. Da die Kosten für Polizei und Gerichtswesen im zweiten Fall aber höher sind, spricht aus Effizienzgründen alles für die erste Alternative, d.h. höhere Strafe bei geringerer Wahrscheinlichkeit, erwischt zu werden.⁶⁷ BECKER selber wandte das ökonomische Modell später auf weitere Rechtsgebiete wie das Familienrecht an, die traditionellerweise der ökonomischen Analyse als nicht zugänglich betrachtet wurden.⁶⁸ Bereits früher hatte er Anti-Diskriminierungs-Gesetze ökonomisch analysiert und war zum Schluss gekommen, dass in einer Wettbewerbswirtschaft der diskriminierende Arbeitgeber den Nachteil der Diskriminierung erleidet.⁶⁹ 1992 erhielt GARY S. BECKER den Nobelpreis dafür, dass er das Anwendungsgebiet der ökonomischen Theorie auf Gebiete ausgeweitet hatte, die früher anderen Sozialwissenschaften wie der Soziologie oder Kriminologie vorbehalten waren.⁷⁰

- 30 RICHARD POSNER schliesslich veröffentlichte 1973 das einflussreichste Lehrbuch der ÖAR. Obwohl bereits bei Erscheinen heftig kritisiert,⁷¹ ist „*Economic Analysis of Law*“ inzwischen in der sechsten Auflage (2002) erschienen und ist eines der fünf meistzitierten juristischen Bücher.⁷² Es zeichnet sich durch eine unerhörte Breite aus: vom Sachenrecht über Wettbewerbsrecht zu Steuerrecht und Verfassungsrecht bis hin zur Rechtsgeschichte lässt es kein Gebiet unberührt von der ökonomischen Analyse. POSNER ist nicht ziemplich, er ist einer der aggressivsten Vertreter der ÖAR. So tritt er beispielsweise für die Versteigerung von Adoptionsrechten an den Meistbietenden ein;⁷³ eine Position, die ihm selbst innerhalb der *Law and Economics* Bewegung Kritik eingetragen hat.⁷⁴ POSNER scheint ein gewisses Vergnügen zu haben an der Provokation durch Sätze wie „Der Wettbewerb unter schwangeren Frauen, ihre elterlichen Rechte zu verkaufen, würde [in einem freiem Markt]

⁶⁷ RICHARD A. POSNER, *Economic Analysis of Law*, 6. Aufl., New York 2002, 226.

⁶⁸ GARY S. BECKER, *Treatise on the Family*, Cambridge 1981.

⁶⁹ GARY S. BECKER, *The Economics of Discrimination*, Chicago 1957.

⁷⁰ Pressemitteilung der Schwedischen Akademie für Wissenschaften vom 13. Oktober 1992, erhältlich unter <http://nobelprize.org/economics/laureates/1992/press.html> (besucht am 30. März 2005).

⁷¹ ARTHUR A. LEFF, *Economic Analysis of Law: Some Realism About Nominalism*, *Virginia Law Review* 1974, 451-482; A. MITCHELL POLINSKY, *Economic Analysis as a Potentially Defective Product: A Buyer's Guide to Posner's *Economic Analysis of Law**, *Harvard Law Review* 1974, 1655-1681.

⁷² FRED R. SHAPIRO, *The Most-Cited Legal Books Published Since 1978*, *Journal of Legal Studies* 2000, 397-406, 404.

⁷³ ELISABETH M. LANDES/RICHARD A. POSNER, *The Economics of the Baby Shortage*, *Journal of Legal Studies* 1978, 323-348; RICHARD A. POSNER, *The Regulation of the Market in Adoption*, *Boston University Law Review* 1987, 59-72.

⁷⁴ GARY S. BECKER/KEVIN M. MURPHY, *The Family and the State*, *Journal of Law and Economics* 1988, 1-18, 14.

den Preis dieser Rechte auf ein Niveau drücken, das nur unwesentlich über den medizinischen und Unterhaltskosten einer Schwangerschaft liegt“.⁷⁵ Seinem Einfluss haben sie jedenfalls nicht geschadet. POSNER und BECKER unterhalten übrigens seit Anfang 2005 gemeinsam ein Internet-Journal („Blog“), in dem sie provokative Thesen (z.B. diejenige, den Drogenhandel zu legalisieren) informell darstellen und mit ihren Lesern diskutieren.⁷⁶

II. Erfolg der Ökonomischen Analyse des Rechts in den USA

- 31 Die Ökonomische Analyse des Rechts ist in den USA zur dominierenden rechtsphilosophischen Strömung geworden.⁷⁷ Selbst für Kritiker der ÖAR ist “das Blühen der Ökonomischen Analyse des Rechts unbestreitbar”,⁷⁸ hat sie “das amerikanische Rechtsdenken transformiert”⁷⁹ und die Effizienzanalyse “unerreichten Erfolg in der Akademie und an den Gerichten” genossen.⁸⁰ Für BRUCE ACKERMAN ist die ÖAR die “wichtigste Entwicklung der Rechtswissenschaft im 20. Jahrhundert”.⁸¹ Der Anteil der Artikel in den *Law Journals* von Chicago, Harvard, Stanford und Yale, die sich der Ökonomischen Analyse des Rechts bedienen, stieg bis 1985 auf rund einen Viertel (siehe Tabelle 1).⁸²

<i>Erscheinungsjahr des Bandes</i>	<i>Anzahl ÖAR-Artikel</i>	<i>Gesamtzahl der Artikel</i>	<i>Prozentualer Anteil ÖAR</i>
1960-61	3	51	6 %
1965-66	5	70	7 %
1970-71	15	53	28 %
1975-76	8	56	14 %
1980-81	18	54	33 %
1985-86	14	59	24 %

Tabelle 1: ÖAR Aufsätze in führenden US Law Journals (aus ELLICKSON, 1989)

- 32 RICHARD POSNER ist inzwischen der meistzitierte Jurist in den USA,⁸³ und dürfte, da der Output an juristischen Arbeiten in keinem Land grösser ist als in den USA, der meistzitierte Jurist der Welt sein.

⁷⁵ RICHARD A. POSNER, *Sex and Reason*, Cambridge 1994, 412 (Paperback-Ausgabe).

⁷⁶ www.becker-posner-blog.com (besucht am 30. März 2005).

⁷⁷ EIDENMÜLLER, FN 56, 19.

⁷⁸ MARC GALANTER/MARK ALAN EDWARDS, Introduction: The Path of the Law Ands, *Wisconsin Law Review* 1997, 375-387, 378.

⁷⁹ GALANTER/EDWARDS, FN 78, 378.

⁸⁰ GALANTER/EDWARDS, FN 78, 381.

⁸¹ COOTER/ULEN, FN 60, 2.

⁸² ROBERT C. ELLICKSON, Bringing Culture and Human Frailty to Rational Actors: A Critique of Classical Law and Economics, *Chicago-Kent Law Review* 1989, 23- 55, 28.

⁸³ FRED R. SHAPIRO, The Most Cited Legal Scholars, *Journal of Legal Studies* 2000, 409- 426, 424. Die Daten von SHAPIRO beruhen auf dem *Social Sciences Citation Index* von 1956-1999. Der SSCI umfasst rund 1'000, vorwiegend US-amerikanische, sozialwissenschaftliche Zeitschriften. Da auch wirtschaftswissenschaftliche, philosophische und soziologische Zeitschriften erfasst werden, reflektiert die Gesamtzahl der Zitate eines Autors auch seinen Einfluss ausserhalb der Rechtswissenschaft.

<i>Name</i>	<i>Zitate</i>
Richard A. Posner (1939-)	7,981
Ronald Dworkin (1931-)	4,488
Oliver Wendell Holmes, Jr. (1841-1935)	3,665
John Hart Ely (1938-2003)	3,032
Roscoe Pound (1870-1964)	3,018

Tabelle 2: Meistzitierte Juristen (aus SHAPIRO, 2000)

- 33 Abbildung 1 ist eine grafische Darstellung der Treffer, die eine Suche in der „*Law Journals*“ Datenbank von *LexisNexis* ergeben hat.⁸⁴ Für jeden der drei Suchstrings [„*economic analysis of law*“], [„*united states supreme court*“ & „*chief justice*“] und [kahneman & tversky] wurde die Suche, beginnend mit 1985, jeweils auf ein Jahr eingeschränkt. Beispielsweise ist in 127 Artikeln, die 1985 erschienen sind, „*economic analysis of law*“ vorgekommen, während in 59 Artikeln „*united states supreme court*“ und „*chief justice*“ vorgekommen sind. Die Suche nach „*united states supreme court*“ und „*chief justice*“ dient dazu, die Resultate zu normalisieren.⁸⁵ Da dieser Suchstring weder mit der ÖAR noch mit der BLE speziell zusammenhängt, zeigt er (i) die Ausweitung der Datenbank (immer mehr Zeitschriften sind elektronisch recherchierbar) und (ii) die steigende Menge akademischer Publikationen. Wie unschwer zu erkennen ist, sind die Zitate für „*economic analysis of law*“ zwischen 1991 und 1998 weniger stark angestiegen als die Menge der elektronisch verfügbaren Artikel. Seit 1999 sinken die Treffer für „*economic analysis of law*“ mehr oder weniger im Gleichschritt mit den Treffern der Kontrollsuche. Die Treffer für [kahneman & tversky]⁸⁶ hingegen steigen seit 1998 gegen den Trend. Dies ist ein quantitativer Indikator für das Interesse, das seit etwa 1998 der *Behavioral Law and Economics* entgegengebracht wird.

⁸⁴ Eine komplette Liste der in der *LexisNexis Law Review* Datenbank erfassten Zeitschriften ist unter www.lexisnexis.com/sourcelists/pdfs/lawrev.pdf erhältlich (besucht am 1. April 2005).

⁸⁵ Eine Suche nur nach „*united states supreme court*“ ergab zu viele Treffer.

⁸⁶ Beide Namen müssen vorkommen, aber nicht in der genauen Reihenfolge.

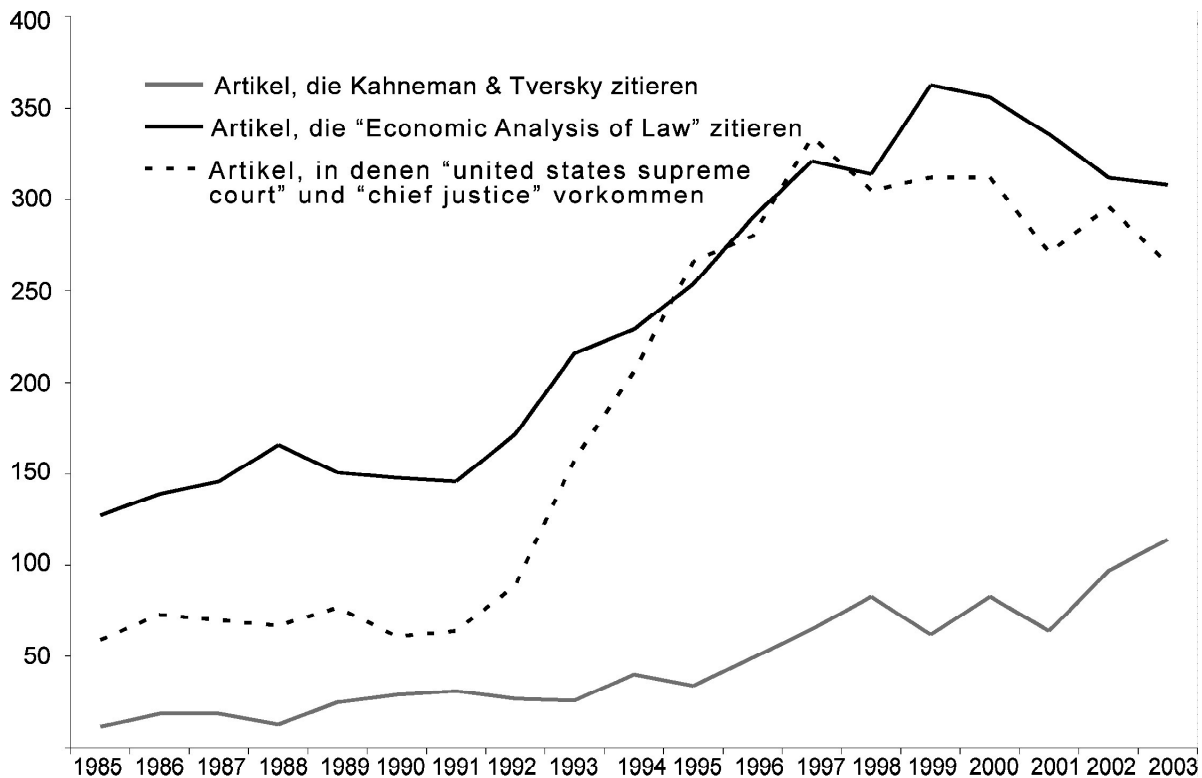


Abbildung 1: Grafische Darstellung der Treffer für verschiedene Suchanfragen in LexisNexis

- 34 Die Ökonomische Analyse des Rechts ist in den USA nicht auf die akademische Diskussion beschränkt geblieben. In zahlreichen Gerichtsurteilen wird mit explizit ökonomischen Überlegungen argumentiert.⁸⁷ Dies ist auch darauf zurückzuführen, dass prominente Vertreter der ÖAR, darunter POSNER, CALABRESI und EASTERBROOK, Richter an Bundesberufungsgerichten sind.⁸⁸ Das instrumentelle Rechtsverständnis der *Legal Realists* führte dazu, dass sich amerikanische Gerichte nicht scheuen, Entscheide ausdrücklich mit rechtspolitischen Argumenten zu begründen und auf Zweckmässigkeitsüberlegungen zu stützen.⁸⁹
- 35 In Europa hat die ÖAR hingegen bis heute kaum Spuren in der Rechtsprechung hinterlassen.⁹⁰ Dies ist auch auf ein unterschiedliches Verständnis der richterlichen Aufgabe zurückzuführen. Während in Kontinentaleuropa die Aufgabe des Richters primär darin besteht, Gesetze zu interpretieren, schafft der Richter in den USA zumindest in den Kerngebieten des *Common Law* das Recht. Weitere Gründe für den bisher geringen Einfluss der ÖAR auf die Rechtsprechung in Europa sind sicherlich das durch die Ausbildung bedingte mangelnde ökonomische Verständnis der Richter,⁹¹ die Anwendung traditioneller

⁸⁷ EIDENMÜLLER, FN 56, 405 ff.

⁸⁸ POSNER ist seit 1981, EASTERBROOK seit 1985 Richter am *7th Circuit Court*, CALABRESI seit 1994 am *2nd Circuit Court*.

⁸⁹ EIDENMÜLLER, FN 56, 409.

⁹⁰ EIDENMÜLLER, FN 56, 19.

⁹¹ GÉRARD HERTIG, Switzerland, *International Review of Law and Economics* 1991, 293-300, 300.

Dogmatik, um die eigene Karriere nicht zu gefährden⁹² und die in Europa von Anfang an heftige Kritik an der ÖAR durch etablierte Rechtslehrer anführen.⁹³

- 36 In der juristischen Literatur wird in Europa eine intensive theoretische Diskussion über die Vor- und Nachteile der ÖAR geführt.⁹⁴ In die Standardwerke der juristischen Methodenlehre ist die ÖAR aber bislang nicht vorgedrungen,⁹⁵ was daran liegen mag, dass sich die deutschsprachige Jurisprudenz in erster Linie als „Rechtsprechungswissenschaft“ versteht und die ÖAR in Europa aus den bereits angetönten Gründen eher eine Gesetzgebungstheorie ist.⁹⁶

III. Gründe für den Erfolg der Ökonomischen Analyse des Rechts

- 37 Die ÖAR kommt mit wenigen Prämissen aus, die eine Vielzahl von verschiedenen Phänomenen erklären können. Die ÖAR ist ein kohärentes System. Kritisieren kann man ihre Prämissen, wenn man aber diese akzeptiert, sind die daraus gezogenen Schlussfolgerungen logisch und widerspruchsfrei. Die ökonomische Analyse ist daher oft elegant, während traditionelle juristische Dogmatik an mittelalterliche Scholastik erinnert. RONALD COASE hat das einmal zugespitzt so formuliert: „[Der Physiker] Ernest Rutherford meinte, dass jede Wissenschaft entweder Physik oder Briefmarkensammeln sei. Damit will er vermutlich sagen, dass sie entweder analysiert oder ein Ablagesystem betreibt. Ein grosser Teil, vielleicht der grösste Teil, der juristische Lehre ist Briefmarkensammeln. *Law and Economics* wird dies wahrscheinlich ändern und hat bereits begonnen, es zu ändern”.⁹⁷
- 38 Soweit die Ökonomische Analyse des Rechts Folgen rechtlicher Regeln auf der Grundlage des ökonomischen Verhaltensmodells prognostiziert, macht sie Aussagen, die grundsätzlich empirisch überprüft werden können. Ökonomische Prognosen sind falsifizierbar. Für die Rechtswissenschaft eröffnet sich dadurch die Möglichkeit, sich zu einer Realwissenschaft im Sinne der modernen Wissenschaftstheorie zu entwickeln.⁹⁸ KAHNEMAN, TVERSKY und anderen Psychologen gelang es nachzuweisen, dass die Prognosen des wirtschaftswissenschaftlichen Verhaltensmodells unter Laborbedingungen und im wirklichen Leben unter ganz bestimmten Bedingungen nicht zutreffen. Sie schlagen deshalb Modifikationen des *homo oeconomicus* vor, die das Modell der Wirklichkeit annähern und daher die Prognosequalität verbessern. *Behavioral Law and Economics* macht sich diese Einsichten für die Analyse des Rechts zu Nutze.

⁹² KIRCHNER, FN 95, 285.

1 ⁹³ KARL-HEINZ FEZER, Aspekte einer Rechtskritik an der economic analysis of law und am property rights approach, Juristen-Zeitung 1986, 817-864.

⁹⁴ EIDENMÜLLER, FN 56, und JANSON, FN 57, bieten einen ausgezeichneten Überblick über die Debatte.

⁹⁵ CHRISTIAN KIRCHNER, The Difficult Reception of Law and Economics in Germany, International Review of Law and Economics 1991, 277-292, 279.

⁹⁶ EIDENMÜLLER, FN 56, 414 ff.

⁹⁷ RONALD COASE, Law and Economics at Chicago, Journal of Law and Economics 1993, 239-254, 254.

⁹⁸ EIDENMÜLLER, FN 56, 6.

- 39 Der normative Massstab der ÖAR, die ökonomische Effizienz, hat den Vorteil der (scheinbaren) Wertfreiheit und Klarheit. „Gerechtigkeit“ dagegen ist ein unklares und umstrittenes normatives Konzept. Bekanntlich kann man sich seit 2'500 Jahren nicht auf eine Definition einigen. Das in der ÖAR meist verwendete Kaldor-/Hicks-Kriterium ist dagegen vergleichsweise klar. Für EIDENMÜLLER besitzt ökonomische Effizienz zudem nach dem Zusammenbruch der kommunistischen Planwirtschaften erhebliche Plausibilität. Die allgemeine Tendenz, neue Strukturen auf der Basis marktwirtschaftlicher Prinzipien zu schaffen, lege es nahe, auch das Recht nach Kriterien der ökonomischen Effizienz zu gestalten.⁹⁹
- 40 *Last but not least* muss man auch darauf hinweisen, dass die Ökonomische Analyse des Rechts oft zum Schluss kommt, dass es am besten ist, wenn sich der Staat nicht einmischt und die freien Kräfte des Marktes walten lässt. Solange Austauschprozesse an Märkten möglich sind, ist das Recht irrelevant, weil sich das gewünschte (effiziente) Ergebnis auf den Märkten von selbst einstellen wird. Aufgabe des Rechts ist es, diese Austauschprozesse nicht zu stören und sie zu erleichtern. Erst wenn Märkte versagen, beispielsweise aufgrund prohibitiv hoher Transaktionskosten, soll das Recht eingreifen und die Lösung rekonstruieren, die sich bei funktionierendem Markt von selbst ergeben hätte. Auch wenn die ÖAR nicht ausschliesslich anti-interventionistisch ist,¹⁰⁰ so ist die allgemeine Stossrichtung doch unübersehbar. Es ist klar, dass die Politik des *laissez-faire*, die Vertreter der ÖAR oft propagieren, Interessen dient, die nicht ganz ohne Einfluss sind;¹⁰¹ das Kaldor/Hicks-Kriterium, wie es in der ÖAR angewendet wird (in der Form von Kosten/Nutzen-Analysen), bevorzugt reiche Individuen, da sie in der Lage sind, mehr für die Güter zu bezahlen, die ihnen wichtig sind.¹⁰²

IV. Rationales Verhalten im Sinne der ökonomischen Theorie

- 41 Kern der Ökonomischen Analyse des Rechts bildet die Annahme, dass der Mensch rational seine Ziele zu verwirklichen sucht.¹⁰³ In dieser Allgemeinheit kann man dieser Aussage kaum widersprechen; in dieser Allgemeinheit taugt sie aber auch nicht, Voraussagen über das konkrete Verhalten von Menschen zu machen. Sie lässt offen, welches die Ziele sind und ebenso undefiniert, welches Verhalten als „rational“ anzusehen ist. Will man daher Prognosen für das menschliche Verhalten treffen, muss man weitere Prämissen postulieren. Mit weiteren Prämissen wird das Modell „dichter“ und erlaubt genauere Voraussagen. „Dünne“ Modelle mit wenigen Prämissen erlauben keine Voraussagen und können daher

⁹⁹ EIDENMÜLLER, FN 56, 6 f.

¹⁰⁰ So EIDENMÜLLER, FN 56, 68.

¹⁰¹ BRIAN LEITER, Holmes, Economics, and Classical Realism, in: STEVEN J. BURTON (Hrsg.), *The Path of the Law and Its Influence: The Legacy of Oliver Wendell Holmes, Jr.*, Cambridge 2000, 285-325, 308: “[...] the ascendancy of economics as a force in public policy has much more to do with its ability to rationalize policies that tangibly benefit ruling groups in American society than with its empirical and cognitive credentials”.

¹⁰² EIDENMÜLLER, FN 56, 246.

¹⁰³ POSNER, FN 67, 3.

auch nicht falsifiziert werden.¹⁰⁴ Zugespißt gesagt: dichte Modelle prognostizieren *ex ante*, dünne Modelle erklären *ex post*. Die verschiedenen Modelle des rationalen Nutzenmaximierers lassen sich nach der Stringenz ihrer Prämissen auf einem Kontinuum anordnen.¹⁰⁵ Die Modelle, die praktisch die grösste Relevanz haben, d.h. von den meisten Vertretern der ÖAR verwendet werden, sind in der Abbildung 2 fett dargestellt.

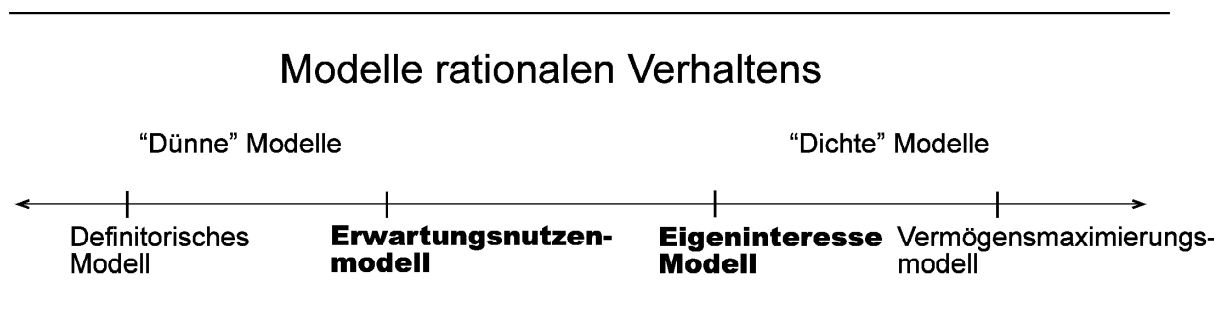


Abbildung 2: Modelle rationalen Verhaltens (nach Korobkin/Ulen, 2000)

A. Definitorisches Modell

42 In seiner dünnsten Version besagt das Modell des rationalen Nutzenmaximierers bloss, dass der Mensch seine Ziele rational zu verwirklichen sucht. Dieser Annahme widerspricht nur, wer gegen seine eigenen Interessen handelt. Praktisch lässt sich das Modell aber nicht falsifizieren, denn sowohl die Mittel wie auch die Ziele des Verhaltens werden durch das beobachtete Verhalten selbst definiert. Weil man annimmt, dass jedes Verhalten rational die Ziele des Handelnden verwirklicht, ist jedes auch noch so seltsame beobachtete Verhalten per definitionem rational. Aus dem beobachteten Verhalten kann man schliessen, welche Ziele (Präferenzen) die Menschen haben und welche Strategien (Mittel) sie verwenden, um diese Ziele zu erreichen. Soll Recht als soziales Steuerungsmittel dienen, muss es aber möglich sein, Voraussagen über das künftige Verhalten der Menschen unter der Geltung verschiedener Normen zu treffen, und dazu ist das dünne Modell rationalen Verhaltens nicht in der Lage. Wie LEFF bemerkt, ersetzt diese Version des *homo oeconomicus* eine normative oder empirische Prämisse durch eine Definition. LEFF spricht daher verächtlich von „*American Legal Nominalism*“.¹⁰⁶

B. Erwartungsnutzenmodell (*expected utility theory*)

43 Das in der modernen Ökonomik am häufigsten verwendete Modell menschlichen
Verhaltens ist das Erwartungsnutzenmodell (*expected utility theory*).¹⁰⁷ Das

¹⁰⁴ Die Terminologie (“*thick*” und “*thin*”) und der Aufbau der Darstellung folgt RUSSEL B. KOROBKIN/THOMAS S. ULEN, Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics, California Law Review 2000, 1053-1144.

¹⁰⁵ Nach KOROBKIN/ULEN, FN 104, 1061.

¹⁰⁶ LEFF, FN 71, 459.

¹⁰⁷ KOROBKIN/ULEN, FN 104, 1062.

Erwartungsnutzenmodell ist dünn in dem Sinn, als es die Präferenzen oder Ziele der Handelnden nicht konkretisiert. Die Präferenzen der Handelnden sind exogen, das Erwartungsnutzenmodell sagt nicht, welchen „Nutzen“ die Handelnden anstreben werden. Es sagt jedoch, dass es bestimmte Mittel gibt, das Ziel zu erreichen. Verletzt der Handelnde diese Mittel, handelt er irrational im Sinne des Erwartungsnutzenmodells.

- 44 Muss der Entscheidende eine riskante Wahl treffen, sagt das Erwartungsnutzenmodell voraus, dass er jedem möglichen Ausgang x der Wahl einen Nutzen u zuschreibt. Wenn der Ausgang sicher ist – die Wahrscheinlichkeit p , dass er eintritt, daher 1 ist – entspricht der erwartete Nutzen dem Nutzen der Alternative. Für eine riskante Wahl muss der Nutzen einer Alternative mit der Wahrscheinlichkeit jedes ihrer möglichen Ausgänge (*outcomes*) multipliziert werden. Der erwartete Nutzen von n sich gegenseitig ausschliessenden *outcomes* x berechnet sich daher nach der Formel

$$\text{Erwarteter Nutzen} = \sum_{i=1}^n p_i \cdot u(x_i) \quad (1)$$

- 45 Der *homo oeconomicus* des Erwartungsnutzenmodells vergleicht den erwarteten Nutzen der sicheren Alternative mit dem erwarteten Nutzen der riskanten Alternative und wählt diejenige Alternative mit dem höheren Nutzen.
- 46 Sieht man von der mathematischen Formulierung des Erwartungsnutzenmodells ab, besagt es nichts anderes, als dass der Handelnde eine Kosten-/Nutzenanalyse durchführt und diejenige Wahl trifft, die den erwarteten Netto-Nutzen (d.h. erwarteter Nutzen abzüglich erwarteter Kosten) maximiert.¹⁰⁸ Ein Handelnder verletzt die Voraussagen des Erwartungsnutzenmodells, wenn er eine Wahl trifft, die seinen Netto-Nutzen nicht maximiert. Da es unmöglich ist, vorauszusagen, welche Wahl optimal ist, ohne die Nutzenfunktion des Handelnden zu kennen und Nutzenfunktionen schwierig (aber nicht unmöglich) zu eruieren sind, können die Voraussagen des Erwartungsnutzenmodells oft nicht direkt falsifiziert werden.¹⁰⁹ Die Präferenzen des Handelnden müssen aber gemäss dem Erwartungsnutzenmodell in seiner Formulierung durch VON NEUMANN und MORGENSTERN, welches wohl das in der Ökonomik am weitesten verbreitete Modell des Erwartungsnutzens ist, gewisse *Axiome* erfüllen, deren wichtigste die Folgenden sind:¹¹⁰

Vergleichbarkeit: der Handelnde ist in der Lage, den Nutzen aller denkbaren Alternativen zu vergleichen.

Transitivität: wenn der Handelnde die Alternative A der Alternative B vorzieht und die Alternative B der Alternative C, dann zieht er die Alternative A der Alternative C vor.

¹⁰⁸ KOROBKIN/ULEN, FN 104, 1063.

¹⁰⁹ Korobkin/Ulen, FN 104, 1064.

¹¹⁰ VON NEUMANN/MORGENSTERN, FN 10. Die formalisierte Theorie von VON NEUMANN und MORGENSTERN ist für Juristen ohne vertiefte mathematische Kenntnisse kaum verständlich. REID HASTIE/ROBYN M. DAWES, *Rational Choice in an Uncertain World*, Thousand Oaks 2001, 256 ff. bieten eine auch für den Laien verständliche Einführung in die *expected utility theory* gemäss VON NEUMANN und MORGENSTERN.

Berechnung der Wahrscheinlichkeiten:

der Handelnde berechnet die Wahrscheinlichkeiten der *outcomes* gemäss den Gesetzen der Wahrscheinlichkeitstheorie.

Unabhängigkeit:

wenn der Handelnde eine Alternative A einer Alternative B vorzieht, muss er auch eine Wahl, entweder Alternative A mit einer gewissen Wahrscheinlichkeit oder Alternative C mit einer gewissen Wahrscheinlichkeit zu erhalten und diese der Wahl, entweder Alternative B mit einer gewissen Wahrscheinlichkeit oder Alternative C mit einer gewissen Wahrscheinlichkeit zu erhalten, vorziehen.¹¹¹ D.h. Elemente, die bei beiden Wahlen identisch sind (die Wahrscheinlichkeit, C zu erhalten), spielen bei der Wahl zwischen A und B keine Rolle.

- 47 Eine weitere Annahme, nämlich dass die Präferenzen des Handelnden relativ konstant sind, ist zwar nicht zwingend für das Erwartungsnutzenmodell, muss aber getroffen werden, wenn man vernünftige Prognosen machen will. Zwar wird nicht bestritten, dass sich Präferenzen langfristig ändern können. Soweit aber Präferenzen nicht unabhängig von den Handlungen der Individuen erfasst werden können, ist es wenig sinnvoll, Veränderungen des menschlichen Verhaltens durch Veränderungen der Präferenzen zu erklären. Jede beliebige Verhaltensänderung kann nachträglich durch geänderte Präferenzen erklärt werden; das Modell verliert dadurch seine Prognosefähigkeit und kann nicht widerlegt werden. Die ÖAR nimmt daher an, dass die Präferenzen konstant sind und beobachtete Verhaltensänderungen auf geänderte Restriktionen (Anreize) zurückzuführen sind.¹¹² Nach BECKER bietet die Annahme konstanter Präferenzen ein stabiles Fundament, um Prognosen zu machen und bewahrt den Analysten davor, offensichtliche Widersprüche zu seinen Voraussagen einfach mit der Veränderung der Präferenzen zu erklären.¹¹³
- 48 Eine weitere Annahme, die ebenfalls nicht zwingend ist, aber häufig getroffen wird, ist diejenige, dass der Grenznutzen zählbarer Güter (und damit insbesondere des Einkommens) abnimmt. Die meisten Glücksspieler, die ein Spiel angeboten bekommen, entweder mit einer Wahrscheinlichkeit von 0,6 Fr. 2 Mio. zu gewinnen oder eine Million auf Nummer sicher, werden die sichere Million vorziehen und auf die Chance, zwei Millionen zu gewinnen, dankend verzichten. Die zweite Million ist offensichtlich weniger wert als die erste Million: der Grenznutzen des Geldes nimmt ab. DANIEL BERNOULLI hat diese Intuition 1738 als erster formalisiert. So wie die Wirkung einer Kerze in einem dunklen Raum grösser ist als bei Sonnenschein, so ist der Nutzen eines Vermögensanstiegs grösser, wenn das Vermögen klein ist. Nach BERNOULLI ist der Nutzen eines Gewinns umgekehrt proportional zum Vermögen, das man bereits besitzt.¹¹⁴ Erwartungsnutzenmodelle nehmen

¹¹¹ Formal: $A \geq B$ ausschliesslich wenn $(A \geq C) \geq (B \geq C)$.

¹¹² KIRCHGÄSSNER, FN 62, 39 f.

¹¹³ GARY S. BECKER, *The Economic Approach to Human Behavior*, Chicago 1976, 5.

¹¹⁴ ELDAR SHAFIR, *Cognitive Psychology of Decision Biases*, in: NEIL J. SMELSER/PAUL B. BALTES, *International Encyclopedia of the Social & Behavioral Sciences*, Amsterdam etc. 2001, 3296-3300, 3297.

daher in der Regel an, dass die Nutzenfunktion u konkav ist, was dazu führt, dass der Handelnde risikoscheu ist.¹¹⁵

- 49 Ob sich Handelnde gemäss den Axiomen des Erwartungsnutzenmodells verhalten, lässt sich empirisch prüfen.¹¹⁶ Die Vertreter der *Behavioral Law and Economics* halten dafür, dass die Axiome des Erwartungsnutzenmodells systematisch verletzt werden, so dass es durch ein empirisch besser belegtes Modell, beispielsweise der *Prospect Theory* von KAHNEMAN und TVERSKY, ersetzt werden muss (hinten, S. 56 ff.).

C. Eigeninteresse

- 50 Dass sich ein Mensch eigennützig verhält bedeutet erst einmal nichts mehr, als dass er seinen Entscheidungen seine *eigene* Nutzenfunktion zu Grunde legt.¹¹⁷ Es besagt aber nichts über den Inhalt des Nutzens. Was dem Individuum Nutzen verschafft, ist eine Frage der inhaltlichen Auffüllung der zweiten Komponente vom „Eigennutzen“; d.h. des Nutzens. Neben monetären können auch andere Vorteile, wie Sozialprestige oder Fortpflanzung, Nutzen bringen. Der Nutzenbegriff ist so weit, dass neben immateriellen Gütern selbst *interdependente* Präferenzen (wie Altruismus, Neid, Missgunst) in die Nutzenfunktion Eingang finden können.¹¹⁸ Wenn ein Individuum sein gesamtes Vermögen für einen guten Zweck stiftet, schliesst der Ökonom, dass der Nutzen des Schenkens für diese Person grösser ist als der Nutzen des Vermögens.
- 51 Ein derart weites Verständnis von Eigennutz führt aber dazu, dass jedes Verhalten erklärt, aber keines mehr prognostiziert werden kann.¹¹⁹ Die meisten Modelle des *homo oeconomicus* postulieren daher zusätzlich zu den Prämissen der Erwartungsnutzentheorie, die die Mittel beschreiben, die ein Individuum verwendet, dass der Mensch in seinem eigenen *Interesse (self-interest)* handelt.¹²⁰ Damit werden neben den Mitteln auch die Ziele einer Beschränkung unterlegt. Im Gegensatz zu dünnen Modellen des rationalen Verhaltens führt die Annahme des Eigeninteresses zu direkt observier- und falsifizierbaren Voraussagen. Dem *homo oeconomicus* wird damit aber auch der Egoismus unterstellt, der für seinen schlechten Ruf verantwortlich ist.

D. Vermögensmaximierung

- 52 Die dichtesten Modelle rationalen Verhaltens postulieren, dass das Ziel des Handelnden die Maximierung seiner finanziellen Position ist. Nach diesen Modellen reagiert der Handelnde ausschliesslich auf monetäre Anreize. Sie liegen den meisten ökonomischen Analysen des Verhaltens juristischer Personen zugrunde.¹²¹

¹¹⁵ Ausführlich zum “Gesetz” des abnehmenden Grenznutzens des Vermögens EIDENMÜLLER, FN 56, 43 ff.

¹¹⁶ KOROBKIN/ULEN, FN 104, 1064.

¹¹⁷ JANSON, FN 57, 31.

¹¹⁸ JANSON, FN 57, 32.

¹¹⁹ JANSON, FN 57, 32.

¹²⁰ KOROBKIN/ULEN, FN 104, 1064.

¹²¹ KOROBKIN/ULEN, FN 104, 1066.

V. Kritik an der Ökonomischen Analyse des Rechts

- 53 Kritiker haben an der ÖAR seit jeher sowohl den normativen Standard (ist Effizienz alleine erstrebenswert?) als auch das Modell rationalen Verhaltens (so verhält sich doch niemand!) kritisiert. Dem Schwerpunkt dieser Arbeit entsprechend wird hier vor allem auf die Kritik am *homo oeconomicus* eingegangen.

A. Kritik an der Effizienz als Rechtsziel

- 54 Das utilitaristische Ziel der Nutzenmaximierung verlangt nach interpersonellen Nutzenvergleichen. Um abwägen zu können, ob ein Zustand der Welt den Nutzen der Gesellschaft maximiert, muss man die Vor- und Nachteile, die den einzelnen Mitgliedern der Gesellschaft aus der Änderung entstehen, gegeneinander abwägen können. Solche interpersonellen Nutzenvergleiche sind aus theoretischen und praktischen Gründen kaum durchzuführen. Sie beruhen immer auf einem Werturteil, nämlich auf der impliziten Annahme, dass die Fähigkeit aller Menschen, Befriedigung oder Nutzen zu empfinden, gleich ist.¹²² Das von der Ökonomischen Analyse des Rechts verwendete Kaldor/Hicks-Kriterium kommt in seiner ursprünglichen Fassung durch KALDOR und HICKS ohne interpersonelle Nutzenvergleiche aus.¹²³ In der Praxis findet das Kaldor/Hicks-Kriterium jedoch immer in der Form von Kosten/Nutzen-Analysen Anwendung; d.h. die mit einer bestimmten Massnahme verbundenen Vor- und Nachteile werden monetär bewertet.¹²⁴ Ist der Saldo der Vor- und Nachteile positiv, dann besteht die Massnahme den Kaldor/Hicks-Test.
- 55 Setzt man den Nutzen, den ein Gut für eine Person hat, mit der Bereitschaft dieser Person, für dieses Gut zu bezahlen, gleich, wie das viele Vertreter der ÖAR implizit oder explizit tun, kann man den (so definierten) Nutzen interpersonell vergleichen. Man wird aber mit dem Problem konfrontiert, dass die Bereitschaft zu zahlen von der Fähigkeit zu zahlen abhängt. Dies anerkennt selbst POSNER, von dem das folgende Beispiel stammt: Ein Wachstumshormon ist sehr selten und weil die Nachfrage gross ist, sehr teuer. Eine arme Familie hat ein Kind, das ohne das Hormon zwergwüchsig bleiben wird. Eine reiche Familie hat ein Kind, das normalwüchsig ist, aber mit dem Hormon einige Zentimeter grösser würde. Die reichen Eltern sind in der Lage, das Hormon zu einem Preis zu kaufen, den die arme Familie nicht bezahlen kann. „In dem Sinn, in dem Wert in diesem Buch [*Economic Analysis of Law*] verstanden wird, ist das Wachstumshormon wertvoller für die reiche Familie, weil der Wert durch die Bereitschaft zu bezahlen gemessen wird; aber das Hormon würde grösseres Glück in den Händen der armen Familie als in den Händen der reichen Familie bewirken.“¹²⁵ POSNER gibt selber zu, dass das Kriterium der Effizienz, verstanden als Allokation von Ressourcen, so dass ihr Wert maximiert wird, Grenzen als ethisches Kriterium für sozial- und wirtschaftspolitische Entscheidungen hat (allerdings wendet er es nach diesen warnenden Worten im Rest des Buches ohne Bedenken an).¹²⁶

¹²² EIDENMÜLLER, FN 56, 190.

¹²³ EIDENMÜLLER, FN 56, 51.

¹²⁴ EIDENMÜLLER, FN 56, 192.

¹²⁵ POSNER, FN 67, 11.

¹²⁶ POSNER, FN 67, 11.

Wenn Bereitschaft zu zahlen mit Nutzen gleichgesetzt wird, hängt der normative Status der ÖAR von der ethischen Rechtfertigung der tatsächlich existierenden Vermögensverteilung ab.¹²⁷ POSNER unternimmt aber nie den Versuch, diese zu rechtfertigen, sondern setzt stillschweigend voraus, dass sie „gerecht“ ist. Wer diese Auffassung nicht teilt, wird auch die Schlussfolgerungen der ÖAR über die effiziente Allokation von Ressourcen nicht teilen.

- 56 Ein weiteres Problem des Utilitarismus und der Ökonomischen Analyse des Rechts ist, dass sie die spezielle Stellung von Grundrechten nicht begründen können. Grundrechte sind in einem System, das die Nutzenmaximierung als oberstes Gut ansieht, nur so lange geschützt, als der Nutzen, sie zu missachten, nicht grösser ist als der Nutzen, sie zu achten. Wenn, so POSNER, eine Person gelyncht wird, deren Schuld mit absoluter Sicherheit feststeht und die nach einem ordentlichen Strafverfahren hingerichtet worden wäre, dann ist der einzige Effekt des Lynchmordes (auf die Reichtumsmaximierung), dass die Kosten für den Prozess gespart werden.¹²⁸ Zwar ist auch der POSNER der Auffassung, dass Lynchmorde in aller Regel nicht gerechtfertigt sind, aber es gibt in einem System, das Nutzen oder Reichtum maximiert, eben keine absolut geschützten „Kernbereiche“ von verfassungsmässigen Rechten, die unantastbar sind.
- 57 EIDENMÜLLER kommt aus diesen und weiteren Gründen nach einer sorgfältigen Analyse zum Schluss, dass „der Versuch, das ökonomische Effizienzziel als alleiniges oder überragendes Ziel der Rechtspolitik zu rechtfertigen, als gescheitert angesehen werden [kann]. Der Gesetzgeber mag sich dazu entschliessen, diesem Ziel sektoral in bestimmten Rechtsgebieten unter Abwägung mit anderen Zielen Rechnung zu tragen. Ein *prima facie*-Vorrang für das ökonomische Effizienzziel lässt sich jedoch nicht begründen“.¹²⁹

B. Kritik am Homo Oeconomicus

- 58 Kritik am Menschenbild der Ökonomik ist so alt wie die Ökonomik. Viele Kritiker übersehen aber, dass es nicht genügt, der Annahme rationalen nutzenmaximierenden Verhaltens die Behauptung entgegenzusetzen, dass sich Menschen *nicht* rational und nutzenmaximierend verhalten. Gerade weil die ökonomische Analyse mit spärlichen Annahmen grosse prognostische Erfolge erzielt hat, sind Ökonomen verständlicherweise zurückhaltend, ihre Prämissen aufzugeben.¹³⁰ Selbst ein Kritiker des traditionellen Modells des rationalen Nutzenmaximierers wie ELLICKSON konzediert, dass die Weigerung von POSNER im Jahr 1975,¹³¹ die Kritik von LEFF am Menschenbild der ÖAR ernst zu nehmen, richtig war.¹³² Die ÖAR des Rechts befand sich damals in voller Blüte und erzielte mit den spärlichen Annahmen des traditionellen Modells grosse intellektuelle Erfolge. Zahlreiche

¹²⁷ SINGER, FN 34, 528.

¹²⁸ RICHARD A. POSNER, *Wealth Maximization Revisited*, Notre Dame Journal of Law, Ethics & Public Policy 1985, 85-105, 101.

¹²⁹ EIDENMÜLLER, FN 56, 321.

¹³⁰ Siehe nur MILTON FRIEDMAN, *The Methodology of Positive Economics*, in: MILTON FRIEDMAN, *Essays in Positive Economics*, Chicago 1953, 3-43.

¹³¹ RICHARD A. POSNER, *The Economic Approach to Law*, Texas Law Review 1975, 757-787, 773 ff.

¹³² ELLICKSON, FN 56, 24.

juristische Phänomene konnten auf gemeinsame Ursachen zurückgeführt und einfacher und einleuchtender als mit traditioneller Dogmatik erklärt werden. Die Kritik am Menschenbild der Ökonomie hingegen steckte in den Kinderschuhen und kam häufig über Appelle, man solle doch die psychologischen und gesellschaftlichen Realitäten vermehrt beachten, nicht hinaus.

- 59 Ende der achtziger Jahre des letzten Jahrhunderts hatte sich dieses Bild jedoch gewandelt: die einfachen und offensichtlichen Anwendungen der Ökonomik im Recht waren gemacht, und neue Erkenntnisse stammten von immer komplexeren mathematischen Modellen, die von Juristen ohne entsprechende Ausbildung nicht verstanden werden konnten. Gleichzeitig stehen die Erkenntnisse, die aus den technischen Modellen gewonnen werden, für viele Juristen in keinem Verhältnis mehr zum theoretischen Aufwand.¹³³ Auf der anderen Seite hat die Kritik am *homo oeconomicus* grosse Fortschritte gemacht. Was in den (frühen) siebziger Jahren Behauptungen waren, konnte sich inzwischen auf 15 bis 20 Jahre empirische Forschung stützen. Der Ertrag, diese Erkenntnisse für die Analyse spezifisch juristischer Fragestellungen einzusetzen, so ELLICKSON, sei erheblich grösser als der Ertrag der weiteren Ausdifferenzierung mathematischer Modelle. Er appellierte daher 1989 an die Vertreter der ÖAR, namentlich an POSNER, der sich bereit erklärt hatte, einen Kommentar zum Artikel von ELLICKSON zu verfassen,¹³⁴ sich diesen neuen Erkenntnissen nicht länger zu verschliessen.¹³⁵

1. Kritik an der Annahme (ausschliesslich) eigeninteressierten Verhaltens

- 60 Die Voraussagen des Verhaltens, die sich aus dem Modell des eigeninteressierten *homo oeconomicus* ergeben, wurden durch neuere empirische Forschung wiederholt widerlegt. Menschen verhalten sich beim Gefangenendilemma, bei Ultimatum- und Diktatorspielen nicht so, wie sie gemäss dem Modell des eigeninteressierten Verhaltens sollten. Beispielsweise sind sie bereit, zur Bestrafung anderer erhebliche Kosten auf sich zu nehmen, ohne einen direkten Vorteil daraus zu erlangen, oder sie kooperieren und stellen das gemeinsame Interesse über das (grössere) Eigeninteresse. Einer der führenden Forscher auf diesem Gebiet ist der an der Universität Zürich lehrende ERNST FEHR.¹³⁶ Die Antwort der Ökonomen scheint bislang zu sein, dass sich die beobachteten Verletzungen des orthodoxen Modells auf Kleinkostensituationen beschränken.¹³⁷ Meines Erachtens haben diese Resultate wichtige Konsequenzen für das Normverständnis des *homo oeconomicus*, die hier aber, wie erwähnt, nicht vertieft werden können.

¹³³ ELLICKSON, FN 56, 33.

¹³⁴ RICHARD A. POSNER, The Future of Law and Economics: A Comment on Ellickson, Chicago-Kent Law Review 1989, 57-70.

¹³⁵ ELLICKSON, FN 56, 35 ff.

¹³⁶ ERNST FEHR/URS FISCHBACHER, The Nature of Human Altruism, Nature 2003, 785-791; ERNST FEHR/URS FISCHBACHER/SIMON GÄCHTER, Strong Reciprocity, Human Cooperation and the Enforcement of Social Norms, Human Nature 2002, 1-25; JOSEPH HENRICH/ROBERT BOYD/SAMUEL BOWLES/COLIN CAMERER/ERNST FEHR/HERBERT GINTIS/RICHARD McELREATH, In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies, American Economic Review 2001, 73-78. Diese und weitere Artikel sind erhältlich unter <http://www.iew.unizh.ch/home/fehr/> (besucht am 1. April 2005).

¹³⁷ KIRCHGÄSSNER, FN 62, 157 ff.

2. Kritik an der Annahme (ausschliesslich) rationalen Verhaltens

a. Eingeschränkte Rationalität (*bounded rationality*)

- 61 Der *homo oeconomicus* des neoklassischen Grundmodells prüft in jedem Einzelfall den Nutzen der ihm zur Verfügung stehenden Handlungsmöglichkeiten in Hinblick auf seine Präferenzen und maximiert entweder bei gegebenem Mitteleinsatz seinen Nutzen oder minimiert den Mitteleinsatz unter Konstanthaltung des Nutzens. Das Individuum besitzt vollständige Informationen, d.h. es kennt sämtliche Handlungsalternativen und ihren Nutzen, so dass die Maximierung „nur“ ein (allerdings rechnerisch aufwändiger) mathematischer Prozess ist. Kosten der Informationsbeschaffung oder Verarbeitung entstehen dabei keine, der *homo oeconomicus* kann blitzschnell beliebige Mengen an Informationen verarbeiten.¹³⁸ Bei Entscheidungen unter Risiko weiss der einzelne nicht sicher, welche genauen Folgen sich aus seiner Handlungsalternative ergeben, sondern kennt nur noch die Wahrscheinlichkeiten, mit der bestimmte Handlungen zu bestimmten *outcomes* führen. Der *homo oeconomicus* maximiert dann nicht mehr den tatsächlichen, sondern den erwarteten Nutzen, also den mit der Eintretenswahrscheinlichkeit für ein bestimmtes Ereignis gewichteten Nutzen.¹³⁹ Dieses Menschenbild wird selbst von Anhängern der Ökonomik als Zerrbild bezeichnet.¹⁴⁰ Es ist offensichtlich, dass es nur wenig mit der Realität zu tun hat und auch zu falschen Voraussagen führen muss, wenn man es als Grundlage für ökonomische Modelle verwendet, die sich nicht mit perfekten Märkten befassen.
- 62 „Eingeschränkte Rationalität“ (*bounded rationality*) bezeichnet Modelle der rationalen Entscheidung, die die kognitiven Grenzen des Entscheidenden berücksichtigen; Grenzen sowohl bezüglich des Wissens wie der rechnerischen Kapazität.¹⁴¹ Die Rationalität des einzelnen ist gemäss Modellen der *bounded rationality* eingeschränkt, weil er erstens nicht mehr alle Handlungsalternativen und –folgen kennt und andererseits diese Daten auch nicht kostenlos und unendlich schnell verarbeiten kann. JANSON und andere weisen zu Recht darauf hin, dass genau genommen nicht die Rationalität, sondern die Kenntnis der Entscheidungsalternativen und ihre Verarbeitung eingeschränkt sind.¹⁴²
- 63 Das Augenmerk richtet sich in diesen Modellen daher naturgemäss auf die Suchstrategien, die der einzelne verwendet, um Informationen über Handlungsalternativen und –folgen zu erwerben. Für HERBERT A. SIMON hat das Individuum gewisse Erwartungen (*aspirations*), die die Alternative erfüllen sollte. Sobald der einzelne eine Alternative gefunden hat, die die Erwartungen erfüllt, wird die Suche abgebrochen und diese Alternative, die nicht notwendigerweise die optimale zu sein braucht, gewählt („*satisficing*“ statt „*optimi-*

¹³⁸ JANSON, FN 57, 39.

¹³⁹ JANSON, FN 57, 39.

¹⁴⁰ KIRCHGÄSSNER, FN 62, 17.

¹⁴¹ HERBERT A. SIMON, Bounded Rationality, in: JOHN EATWELL/MURRAY MILGATE/PETER NEWMAN, The New Palgrave: a dictionary of economics, London 1987, Bd. 1, 266-267, 266.

¹⁴² JANSON, FN 57, 40.

zing“).¹⁴³ Wird binnen vernünftiger Zeit keine akzeptable Alternative gefunden, wird das Anspruchsniveau gesenkt, so dass die Suche nicht ewig dauert. SIMON hat sein Modell eingeschränkt rationalen Verhaltens bereits 1955 formalisiert und erhielt 1978 unter anderem dafür den Nobelpreis für Wirtschaft.¹⁴⁴

- 64 Die Institutionenökonomik hat sich das Modell der eingeschränkten Rationalität zu Eigen gemacht, indem sie der Informationssuche und –verarbeitung positive Kosten zumisst. Die eingeschränkte Rationalität wird zur Voraussetzung der Betrachtung, da viele Institutionen (im ökonomischen Sinn)¹⁴⁵ gerade dazu dienen, besser mit ihr umzugehen.¹⁴⁶ Mit anderen Worten werden die Kosten der Informationssuche und –verarbeitung als *Restriktionen* betrachtet, und die Strategien des einzelnen, mit diesen Restriktionen umzugehen – beispielsweise, indem man sich mit einer befriedigenden statt optimalen Entscheidung begnügt – als rationale Reaktion darauf. Ökonomen können daher ohne grosse Probleme behaupten, dass „eine solche auf unvollständiger Information und Informationsverarbeitungskapazität basierende eingeschränkte Rationalität [...] nach alldem jedenfalls ein Unterfall des modernen ökonomischen Rationalitätsbegriffs [ist und bleibt] und [...] allenfalls im Widerspruch zum Modell des situativ-nutzenmaximierenden *homo oeconomicus* der Neoklassik [steht]“. ¹⁴⁷ Im übrigen wird oft angenommen, dass die Strategien der eingeschränkten Rationalität auf Märkten kaum eingesetzt werden, da dort die mit ihnen verbundenen Kosten wegen des Konkurrenzdruckes besonders hoch seien; m. a. W. eine suboptimale Entscheidungsstrategie sofort „bestraft“ wird.¹⁴⁸
- 65 Im Sinne ihres Erfinders ist diese Interpretation der *bounded rationality* nicht. SIMON weist darauf hin, dass die Annahme der Institutionenökonomik, der *homo oeconomicus* optimiere seine Suchstrategie unter Berücksichtigung der Grenzkosten und –nutzen zusätzlicher Informationssuche, die Komplexität der supponierten Berechnung erhöht und nicht vermindert: nun muss der einzelne nicht nur Angebots- und Nachfragekurven berechnen, sondern auch die Vor- und Nachteile, diese Kurven genauer zu berechnen.¹⁴⁹ Trotzdem muss man konstatieren, dass das Konzept der eingeschränkten Rationalität den Rationalitätsbegriff, wie er in der Ökonomischen Analyse des Rechts verwendet wird, nicht ernsthaft zu erschüttern vermochte. Modelle der eingeschränkten Rationalität enthalten alle Elemente, die für das ökonomische Verhaltensmodell entscheidend sind: die Trennung zwischen Präferenzen und Restriktionen, die Bewertung (eines Teils) der Alternativen, die

¹⁴³ HERBERT A. SIMON, Rational Decision-Making in Business Organizations, Economic Sciences 1978, 343-371, 356; erhältlich unter <http://nobelprize.org/economics/laureates/1978/simon-lecture.pdf> (besucht am 7. April 2005).

¹⁴⁴ HERBERT A. SIMON, A Behavioral Model of Rational Choice, Quarterly Journal of Economics 1955, 99-118. Den Nobelpreis erhielt SIMON “for his pioneering research into the decision-making process within economic organizations”.

¹⁴⁵ Institutionen im Sinne der Institutionenökonomik sind vereinfacht gesagt Systeme von Regeln oder Normen (einschließlich der Mechanismen ihrer Durchsetzung).

¹⁴⁶ JANSON, FN 57, 42.

¹⁴⁷ JANSON, FN 57, 42.

¹⁴⁸ KIRCHGÄSSNER, FN 62, 32.

¹⁴⁹ SIMON, FN 143, 359.

Auswahl der (bewerteten) Alternative nach dem relativen Vorteil und damit die systematische Beeinflussbarkeit dieses Verhaltens durch veränderte Umweltbedingungen.¹⁵⁰

b. Heuristics and biases

- 66 Drei wichtige empirische, methodologische und theoretische Stränge der Kritik am Menschenbild des rationalen Nutzenmaximierers, die KAHNEMAN und TVERSKY inspiriert haben, sind die Arbeiten von PAUL A. MEEHL, WARD EDWARDS und diejenigen des eben erwähnten HERBERT A. SIMON.¹⁵¹
- 67 PAUL A. MEEHL veröffentlichte 1954 eine empirische Arbeit, die die Voraussagen von klinisch tätigen Ärzten mit den Voraussagen verglich, die mit einfachen statistischen Modellen gemacht wurden.¹⁵² Er stellte fest, dass die Voraussagen der Experten oft falsch, sie aber nichtsdestotrotz von ihrer Leistung überzeugt waren.¹⁵³ Der Gegensatz von bescheidener Leistung und robustem Selbstbewusstsein legte es nahe, nach kognitiven Mechanismen zu suchen, die zu einleuchtenden, aber falschen Schlüssen führen.
- 68 WARD EDWARDS' wichtiger methodologischer Beitrag war, dass er mit dem Bayes-Theorem einen normativen Standard in die psychologische Forschung einführte, an dem die Rationalität alltäglichen statistischen Denkens gemessen werden konnte.¹⁵⁴ Es war schnell klar, dass intuitives Denken nicht dem „idealen“ normativen Standard entsprach.¹⁵⁵
- 69 Der wichtigste theoretische Beitrag schliesslich stammte von SIMON. Er wies nach, dass die Berechnungen, die gemäss dem traditionellen Modell vom rationalen Menschen verlangt wurden, ein unrealistischer Standard für menschliches Denken sind. Er schlug stattdessen vor, einen realistischeren Standard anzuwenden, der die Grenzen menschlicher Informationsverarbeitung berücksichtigte. Bereits 1957 diskutierte SIMON vereinfachende Strategien, die Menschen anwenden können, um die Komplexität von Entscheidungen zu reduzieren.¹⁵⁶ Inspiriert durch die Beispiele von fehlerhaften Entscheidungen im richtigen Leben, von denen MEEHL berichtete und ausgestattet mit einem normativen Standard, an dem die Korrektheit einer Entscheidung gemessen werden konnten, entwarfen KAHNEMAN und TVERSKY ihr eigenes Modell eingeschränkter Rationalität.¹⁵⁷

¹⁵⁰ KIRCHGÄSSNER, FN 62, 31.

¹⁵¹ THOMAS GILOVICH/DALE GRIFFIN, Introduction - Heuristics and Biases: Then and Now, in: GILOVICH/GRIFFIN/KAHNEMAN (Hrsg.), 1-18, 2. Der Überblick über die Geschichte des *heuristic and biases* Ansatzes folgt weitgehend der erwähnten Einführung von GILOVICH/GRIFFIN.

¹⁵² PAUL A. MEEHL, Clinical versus Statistical Prediction, Minneapolis 1954.

¹⁵³ Eine Übersicht über den aktuellen Forschungsstand bieten ROBYN M. DAWES/DAVID FAUST/PAUL E. MEEHL, Clinical versus Actuarial Judgment, in: GILOVICH/GRIFFIN/KAHNEMAN (Hrsg.), 716-729.

¹⁵⁴ WARD EDWARDS/HAROLD LINDMAN/LEONHARD J. SAVAGE, Bayesian Statistical Inference for Psychological Research, Psychological Review 1963, 193-242.

¹⁵⁵ GILOVICH/GRIFFIN, FN 151, 2.

¹⁵⁶ HERBERT A. SIMON, Models of Man: Social and Rational, New York 1957.

¹⁵⁷ DANIEL KAHNEMAN/PAUL SLOVIC/AMOS TVERSKY, Preface, in: KAHNEMAN/SLOVIC/TVERSKY (Hrsg.), XI-XIII, XI.

- 70 Die dem Forschungsansatz von KAHNEMAN und TVERSKY zu Grunde liegende Idee ist, dass der Mensch eine begrenzte Anzahl von Faustregeln verwendet, um Entscheidungen unter Unsicherheit zu treffen. „[...] Menschen verlassen sich auf eine begrenzte Anzahl von heuristischen Prinzipien, die die komplexe Aufgabe, Wahrscheinlichkeiten zu schätzen und Werte vorauszusagen, vereinfachen. Im Allgemeinen sind diese Heuristiken recht nützlich, aber manchmal führen sie zu schweren und systematischen Fehlern“.¹⁵⁸ In einem Aufsatz mit dem Titel „*Judgment under uncertainty: Heuristics and biases*“, der 1974 in *Science* erschien und inzwischen zum meistzitierten Aufsatz der psychologischen Forschung zu Entscheidungen unter Unsicherheit wurde,¹⁵⁹ hatten KAHNEMAN und TVERSKY ihre Forschungen bis zu diesem Zeitpunkt zusammengefasst und drei Urteilsheuristiken identifiziert, die zu einer Vielzahl von Fehlern (*biases*) führen: die Repräsentativität (*representativeness*, S. 128 ff.), Verfügbarkeit (*availability*) und Verankerung (*anchoring*, S. 77 ff.). (In der juristischen Literatur wird oft nicht zwischen den Strategien (*heuristics*) und den daraus resultierenden Fehlern (*biases*) unterschieden und statt dessen einfach von „*heuristics and biases*“ gesprochen.) Wer beispielsweise gebeten wird, die relative Häufigkeit von Richtern, die Alkoholiker sind, zu schätzen, überlegt sich nach diesem Modell, wie einfach es ist, sich an Berichte über Richter mit Alkoholproblemen zu erinnern (Verfügbarkeit). Geht es darum, die Wahrscheinlichkeit einzuschätzen, dass ein spezifischer Richter Alkoholiker ist, verlässt man sich auf die Ähnlichkeit dieses Richters mit dem Bild eines typischen Alkoholikers (Repräsentativität). In beiden Fällen schliesslich lässt man sich oft von einem Ausgangswert (beispielsweise der Rate der Alkoholiker in der Gesamtbevölkerung) beeinflussen und passt diesen Ausgangswert dann nach oben oder unten an (Verankerung und Anpassung).
- 71 Drei wichtige Merkmale dieses Ansatzes sind, dass die Urteilsheuristiken zwar anhand von Fehlern identifiziert und erforscht wurden, aber an und für sich durchaus vernünftige und keineswegs „irrationale“ Verfahren sind. Zweitens, obwohl Urteilsheuristiken schnell sind, basieren sie auf kognitiven Prozessen, die keineswegs banal sind. Drittens sind die Urteilsheuristiken keine Reaktion auf exzessive Komplexität oder ein Überangebot von Informationen, sondern normale intuitive Reaktionen selbst auf einfachste Fragen zu Wahrscheinlichkeit, Häufigkeit und Voraussagen.¹⁶⁰ Ein häufiges Missverständnis, das aus der Beschreibung der Urteilsheuristiken als „Strategien zur Reduktion der Komplexität von Entscheidungen unter Unsicherheit“ resultiert, ist, dass Menschen die Heuristiken *bewusst* einsetzen würden, um eine schwierige Entscheidung zu vereinfachen. Heuristiken wären dann das Produkt von Faulheit, was impliziert, dass sie nicht verwendet (und die mit ihnen verbundenen Fehler vermieden) werden, wenn sich der Entscheidende nur wirklich bemüht. Die von KAHNEMAN und TVERSKY identifizierten Heuristiken wirken aber unbewusst und daher selbst dann, wenn der Entscheidende allen Grund hat, einen korrekten

¹⁵⁸ TVERSKY/KAHNEMAN, FN 8, 1024.

¹⁵⁹ RÜDIGER F. POHL, Introduction: Cognitive Illusions, in: RÜDIGER F. POHL (Hrsg.), *Cognitive Illusions*, Hove 2004, 1-20, 8.

¹⁶⁰ GILOVICH/GRIFFIN, FN 151, 3.

Entscheid zu treffen, weil sein Entscheid beispielsweise finanzielle Konsequenzen für ihn hat.¹⁶¹

- 72 Der *heuristic and biases* Ansatz war (und ist) ausserordentlich erfolgreich und führte zu einer wahren Flut von Publikationen (wichtige Arbeiten aus den siebziger Jahren sind in der von KAHNEMAN, SLOVIC und TVERSKY herausgegebenen Aufsatzsammlung „*Judgment under uncertainty: Heuristics and biases*“ enthalten, neuere Arbeiten im 2002 erschienenen Sammelband „*Heuristics and Biases: The Psychology of Intuitive Judgment*“ von GILOVICH, GRIFFIN und KAHNEMAN). Die Gründe für den Erfolg sind zahlreich. Einerseits hat der *heuristics and biases* Ansatz sicher vom „Siegeszug“ der Ökonomik profitiert. In dem Masse, in dem in immer mehr Gesellschaftswissenschaften – wie Soziologie, Recht und Politologie – das Menschenbild des *homo oeconomicus* zur Analyse gesellschaftlicher Phänomene herangezogen wurde, wurde auch die (fundierte) Kritik an diesem Menschenbild für immer mehr Gesellschaftswissenschaftler relevant. Die Forschungen von KAHNEMAN und TVERSKY boten einen Ansatzpunkt, das Menschenbild der Ökonomik, das viele Sozialwissenschaftler ablehn(t)en, anzugreifen. In der Ökonomie selber führte (unter anderem) der *heuristic and biases* Ansatz zu den neuen Teilgebieten der *behavioral economics* und der *behavioral finance* (letztere in den neunziger Jahren zunehmend „trendy“ – was Wunder, rationales Verhalten konnte den Börsenboom Ende der Neunziger auf jeden Fall nicht erklären).¹⁶² RICHARD THALER begann bereits in den frühen achtziger Jahren, das Programm von KAHNEMAN und TVERSKY fruchtbar auf ökonomische Fragestellungen anzuwenden.¹⁶³
- 73 Zudem bediente sich das Programm von KAHNEMAN und TVERSKY populärer und einleuchtender Metaphern, die mit der „kognitiven Wende“ der sechziger Jahre eingeführt worden waren. Das Bild des Gehirns als Computer mit limitierten Rechen- und Speicherkapazitäten bot sich geradezu an, durch das Bild von Unterrouتين, die Rechenaufwand sparen, ergänzt zu werden.¹⁶⁴ Auch ist das Forschungsprogramm für Laien ohne Psychologiestudium zugänglich und verständlich. Es verzichtet auf komplexe Modelle und beschränkt sich darauf, Urteilsheuristiken verbal und leicht verständlich zu beschreiben (was ihm innerhalb der Psychologie wiederum Kritik eingetragen hat). Die Experimente, die das empirische Fundament lieferten, gleichen „mehr Cocktail Party Anekdoten als traditionellen Studien der kognitiven Psychologie“.¹⁶⁵ Das macht es einfach, sie weiter zu erzählen, in Vorlesungen, Vorträge und Aufsätze einzubauen, sich an sie zu erinnern und sie zu replizieren – beispielsweise durch einen Doktoranden der Rechtswissenschaft.

¹⁶¹ COLIN F. CAMERER/ROBYN M. HOGARTH, The Effects of Financial Incentives in Experiments: A Review and Capital-Labor Production Framework, *Journal of Risk and Uncertainty* 1999, 7-42.

¹⁶² LAIBSON/ZECKHAUSER, FN 13.

¹⁶³ RICHARD H. THALER, Toward a Positive Theory of Consumer Choice, *Journal of Economic Behavior and Organization* 1980, 39-60.

¹⁶⁴ GILOVICH/GRIFFIN, FN 151, 6.

¹⁶⁵ GILOVICH/GRIFFIN, FN 151, 7.

c. Referenzpunkte, *Framing* und Risikoverhalten

- 74 Ebenfalls auf KAHNEMAN und TVERSKY zurück geht die Forschung zum Risikoverhalten bei der Wahl zwischen Gewinnen und Verlusten. Dabei handelt es sich nicht um ein Phänomen der Informationsverarbeitung, wie die Urteilsheuristiken, sondern eher um die Feststellung, dass Menschen auf die unterschiedliche Darstellung der gleichen Situation unterschiedlich reagieren. *Framing* Effekte sind weniger relevant wegen ihres Beitrages zur Psychologie als wegen ihrer Prognose von Entscheidungen in der Wirklichkeit und der Herausforderung, die sie für das Erwartungsnutzenmodell darstellen.¹⁶⁶
- 75 Die Erwartungsnutzentheorie nimmt Vermögenszustände als Träger des Nutzens an, nicht Vermögensänderungen. Wenn man ein Spiel (*gamble*) evaluiert, bei dem man mit einer Wahrscheinlichkeit von p_1 eine Summe y_1 gewinnen und mit einer Wahrscheinlichkeit von p_2 eine Summe y_2 verlieren kann, so vergleicht man gemäss Erwartungsnutzenmodell den Nutzen von $p_1 \cdot u(w + y_1)$ mit dem Nutzen von $p_2 \cdot u(w - y_2)$ (w steht für *wealth*). Wenn das Nettovermögen des Spielers Fr. 50'000 ist und er ein Spiel angeboten erhält, bei dem er mit einer Wahrscheinlichkeit von 0,1 Fr. 1'000 gewinnen und mit einer Wahrscheinlichkeit von 0,9 Fr. 10 verlieren kann, so vergleicht er nach dem Erwartungsnutzenmodell den Nutzen von $0,1 \cdot u(\text{Fr. } 51'000)$ mit dem Nutzen von $0,9 \cdot u(\text{Fr. } 49'990)$. KAHNEMAN und TVERSKY wunderten sich, warum in allen Spielen, die Versuchspersonen angeboten wurden, die Ergebnisse als Gewinne oder Verluste beschrieben wurden – theoretisch sollte es darauf ja nicht ankommen, sondern nur auf den Endzustand des Vermögens.¹⁶⁷ KAHNEMAN und TVERSKY stellten empirisch fest, dass sich Menschen im Bereich potentieller Gewinne risikoscheu, im Bereich potentieller Verluste aber risikogeneigt verhielten. Auch wird der negative „Nutzen“ (Disutilität) eines Verlustes doppelt so stark empfunden wie der Nutzen eines identischen Gewinnes („*losses loom larger than gains*“). Die Nutzenfunktion ändert sich daher am Referenzpunkt, der Gewinne und Verluste unterscheidet. Weiter stellten KAHNEMAN und TVERSKY fest, dass bei der Evaluation von Spielen die Wahrscheinlichkeiten nicht linear gewichtet werden; ein Unterschied zwischen Sicherheit (Wahrscheinlichkeit = 0) und einer Wahrscheinlichkeit von 0,01 wird viel stärker gewichtet als der (mathematisch identische) Unterschied zwischen einer Wahrscheinlichkeit von 0,4 und 0,41 („*certainty effect*“). 1979 veröffentlichten KAHNEMAN und TVERSKY in der angesehenen ökonomischen Zeitschrift *Econometrica* einen Aufsatz, der die von ihnen empirisch festgestellten Abweichungen von der Erwartungsnutzentheorie in einem formalisierten Modell beschrieb.¹⁶⁸ Die so genannte *Prospect Theory* folgte weitgehend den Vorgaben der rationalen Entscheidungstheorie und modifizierte diese nur insoweit, als notwendig war, um die experimentellen Resultate zu erklären.¹⁶⁹ Die *Prospect Theory* war anfänglich auf die Analyse von einfachen monetären Spielen mit zwei *outcomes* beschränkt. Da sie aber formalisiert war, konnten Ökonomen sie in ihren Modellen benutzen,

¹⁶⁶ DANIEL KAHNEMAN, Preface, in: KAHNEMAN/TVERSKY (Hrsg.), IX-XVII, XV.

¹⁶⁷ KAHNEMAN, FN 166, IX.

¹⁶⁸ KAHNEMAN/TVERSKY, FN 14.

¹⁶⁹ KAHNEMAN, FN 166, X.

was ihren Einfluss stark erhöhte (eine detailliertere Darstellung der Prospect Theory folgt auf S. 53 ff.).¹⁷⁰

- 76 In einem weiteren Artikel beschrieben KAHNEMAN und TVERSKY 1981 *Framing* Effekte.¹⁷¹ Die Präferenz für eine Wahl hängt demgemäss von ihrer *Darstellung* ab, die einigermassen willkürlich ist, da das gleiche Problem in unterschiedlicher Art dargestellt werden kann (siehe S. 54 f. für ein Beispiel). *Framing* ist dabei ein doppeldeutiger Begriff: er bezeichnet einerseits die (vorgegebene) Darstellung des Problems, andererseits aber auch die Interpretation des Problems durch den Entscheidenden selbst.¹⁷² Letzteres lässt sich zugegebenermassen kaum mit Darstellung übersetzen, aber andere deutsche Begriffe wie „Rahmung“ scheinen mir noch weniger elegant und lösen das Problem auch nicht.¹⁷³ Einen Überblick über Klassiker und aktuelle Forschung zu Bewertung von Alternativen unter Berücksichtigung von nichtlinearer Gewichtung der Wahrscheinlichkeit und *Framing*-Effekten bietet der von KAHNEMAN und TVERSKY herausgegebene Sammelband „*Choices, Values and Frames*“ (Cambridge 2000).

d. Kontext und Präferenzen

- 77 Der *homo oeconomicus* des Erwartungsnutzenmodells hat alle Wahlmöglichkeiten (Optionen) nach ihrer Präferenz geordnet hat und wählt – wenn er die Wahl zwischen mehreren Optionen hat – diejenige der erhältlichen Optionen, die in seiner „Präferenz-Rangliste“ zuoberst steht. Optionen, die nicht gewählt werden, sind nach diesem Modell irrelevant,¹⁷⁴ was auch intuitiv einleuchtet: wer lieber eine Kugel Vanille-Eis für Fr. 5 als eine Kugel Erdbeer-Eis für Fr. 6 hat, sollte seine Vorliebe nicht ändern, wenn er erfährt, dass zusätzlich eine Kugel Schokolade-Eis für Fr. 7 erhältlich ist, die er nicht wählt.
- 78 Selbst diese scheinbar unproblematische Annahme des Erwartungsnutzenmodells wird aber regelmässig verletzt. Die Präferenz für eine Option hängt auch vom Kontext ab, in dem die Option präsentiert wird. Die mittlere Option (Kompromiss-Lösung) wird beispielsweise signifikant häufiger gewählt als eine extreme Option.¹⁷⁵ In dem obigen Beispiel würden mehr Leute das Erdbeer-Eis wählen, wenn das teurere Schokolade-Eis ebenfalls auf der Speisekarte ist. Präferenzen sind daher nicht konstant, sondern werden bis zu einem gewissen Grad „*on the spot*“ konstruiert. Ein Beispiel, wie das *Choice-set*, die zur Auswahl stehenden Optionen, auch juristische Entscheidungen beeinflussen kann, wird auf S. 243 ff. besprochen.

¹⁷⁰ ROBERT MACCOUN, Why a Psychologist Won the Nobel Prize in Economics, American Psychological Society Observer 2002, 1-8, 4.

¹⁷¹ AMOS TVERSKY/DANIEL KAHNEMAN, The Framing of Decisions and the Psychology of Choice, Science 1981, 453-458.

¹⁷² KAHNEMAN, FN 166, XIV.

¹⁷³ Bei JOHN R. ANDERSON, Kognitive Psychologie, 2. Auflage, Heidelberg 1996, 335, wird *framing* beispielsweise als „Rahmungseffekt“ übersetzt

¹⁷⁴ AMOS TVERSKY/ITAMAR SIMONSON, Context-Dependent Preferences, Management Science 1993, 117-185; gekürzter Nachdruck in: KAHNEMAN/TVERSKY (Hrsg.), 518-527, 518. Die Seitenzahlen werden gemäss Nachdruck zitiert.

¹⁷⁵ ITAMAR SIMONSON, Choice Based on Reason: The Case of Attraction and Compromise Effects, Journal of Consumer Research 1989, 158-174.

e. Weitere „Kognitive Täuschungen“

- 79 Es gibt keine allgemein anerkannte Definition des Begriffs „kognitive Täuschungen“. Obwohl der Begriff „Kognition“ unumstritten auch die Wahrnehmung umfasst, wird unter dem Begriff „kognitive Täuschungen“, wie er üblicherweise gebraucht wird, ein Sammelbegriff von Täuschungen des Denkens, Urteilens und Erinnerns zusammengefasst, explizit aber keine Täuschungen der Wahrnehmung.¹⁷⁶ Der Begriff „Täuschung“ impliziert eine Abweichung von der „Realität“. Bei Täuschungen der Wahrnehmung und der Erinnerung ist meist unbestritten, wann eine Abweichung vorliegt; die tatsächliche Länge der Linien bei der Müller-Lyer Illusion (Abbildung 3) lässt sich messen und über die Vergangenheit liegen (zumindest in psychologischen Experimenten) Aufzeichnungen vor, mit denen sich die Erinnerungen vergleichen lassen. Bei Täuschungen des Denkens hingegen ist oft bereits der normative Standard umstritten, mit denen der Denkvorgang verglichen werden soll; ohne Standard kann man auch nicht von Abweichung oder Täuschung sprechen.¹⁷⁷

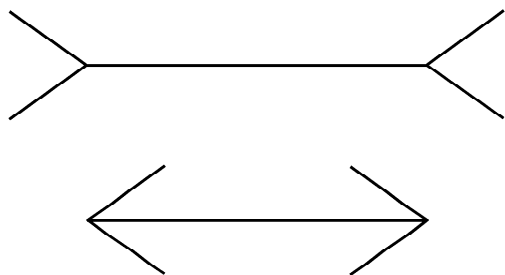


Abbildung 3: Beispiel für eine Illusion der Wahrnehmung: Müller-Lyer Illusion

- 80 Eine Täuschung ist aber mehr als eine bloße Abweichung von der Realität. POHL versucht, folgende Elemente des Begriffs „kognitive Täuschung“ herauszuarbeiten:¹⁷⁸ das Phänomen tritt systematisch, und damit vorhersagbar, auf. Die Fehlleistung erfolgt unbewusst, und Personen, die der Täuschung unterliegen, sind überzeugt, richtig entschieden zu haben. Kognitive Täuschungen sind daher schwierig zu vermeiden. Schliesslich sind die Phänomene deutlich verschieden (*distinct*) von der normalen Informationsverarbeitung. Schlichtes Vergessen oder Missverständnisse zählen daher nicht zu den kognitiven Täuschungen, obwohl sie natürlich auch Fehler sind. Als letzte Eigenschaft von kognitiven Täuschungen könnte man anfügen, dass die meisten Menschen ein „Aha“ Erlebnis haben, wenn ihnen die Täuschung und die richtige Lösung erklärt wird; d.h. sie verstehen und akzeptieren, wo der Fehler lag. Dass es Bedingungen gibt, unter denen die Täuschung verschwindet, ist daher nicht nur für optische (Abbildung 4) sondern auch für kognitive Täuschungen typisch.¹⁷⁹

¹⁷⁶ HELL, FN 11, 318.

¹⁷⁷ POHL, FN 159, 2.

¹⁷⁸ POHL, FN 159, 2 f. Für eine weitere Umschreibung siehe WARD EDWARDS/DETLOF VON WINTERFELDT, Cognitive Illusions and their Implications for the Law, Southern California Law Review 1986, 225-276, 227.

¹⁷⁹ DANIEL KAHNEMAN/AMOS TVERSKY, On the Reality of Cognitive Illusions, Psychological Review 1996, 582-591, 586

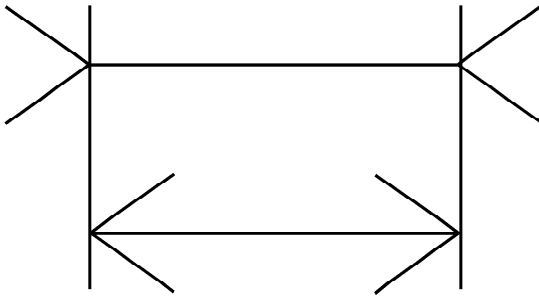


Abbildung 4: „Aufgelöste“ Müller-Lyer Illusion

- 81 Üblicherweise werden die erstmals von KAHNEMAN und TVERSKY beschriebenen *heuristics and biases* zu den kognitiven Täuschungen gezählt,¹⁸⁰ der Rückschaufehler (S. 209 ff.), Gedächtnistäuschungen durch nachträglich eingeführte Informationen, Bestätigungsfehler (S. 178 ff), Selbstüberschätzung und Overconfidence (S. 262 ff.) sowie Attributionsfehler.¹⁸¹ Der Hofeffekt (S. 230 ff.), *Framing*-Effekte sowie Kontrast- und Kompromisseffekte (S. 243 ff.) zählen dagegen traditionellerweise nicht zu den kognitiven Täuschungen. Die beiden Letzteren stellen jedoch wichtige Abweichungen von der Rationalität des *homo oeconomicus* im Sinne der Erwartungsnutzentheorie dar, weshalb es gerechtfertigt scheint, sie hier ebenfalls zu besprechen. Ob der Titel dieser Arbeit deshalb besser „Abweichungen vom rationalen Verhalten im Sinne des Rationalitätsbegriffs der Institutionenökonomik und ihre Auswirkungen auf die Entscheidungen von Richtern und Parteien in Zivil- und Strafverfahren“ heissen müsste, bleibe dahingestellt; „Kognitive Täuschungen vor Gericht“ ist auf jeden Fall kürzer und prägnanter.

f. Zusammenfassung

- 82 Zusammengefasst stellt die psychologische Forschung das Modell des *homo oeconomicus* durch die folgenden empirischen Beobachtungen in Frage:
1. Menschen machen systematische Fehler bei der *Berechnung* von Wahrscheinlichkeiten.
 2. Menschen *gewichten* geringe Wahrscheinlichkeiten zu stark und mittlere bis hohe Wahrscheinlichkeiten zu wenig.
 3. Menschen empfinden Nutzen nicht als Funktion ihres Vermögenszustandes, sondern als Funktion der Veränderung ihres Vermögens von einem *Referenzpunkt* aus, wobei sie sich gegenüber potentiellen Gewinnen risikoscheu, gegenüber potentiellen Verlusten risikogeneigt zeigen.
 4. Menschen empfinden die „Disutilität“ eines Verlustes rund doppelt so stark wie den Nutzen eines gleich grossen Gewinnes.

¹⁸⁰ Obwohl KAHNEMAN und TVERSKY selber den Begriff selten verwenden; der in FN 179 zitierte Aufsatz ist eine Ausnahme, weil er auf einen Aufsatz von GERD GIGERENZER mit dem Titel *How to Make Cognitive Illusions Disappear: Beyond "Heuristics and Biases"* (European Review of Social Psychology 1991, 83-115) reagiert.

¹⁸¹ Siehe die Themen der Aufsätze in HELL/FIEDLER/GIGERENZER (Hrsg.), FN 11, und POHL (Hrsg.), FN 159.

5. Präferenzen sind nicht konstant, sondern hängen vom *Kontext* ab, in dem eine Option präsentiert wird.
- 83 Diese Aufzählung ist nicht abschliessend (zu nennen wäre beispielsweise auch die festgestellte Unfähigkeit, den erfahrenen Nutzen korrekt zu voraussagen),¹⁸² aber diese Arbeit beschränkt sich darauf, die Implikationen der erwähnten Abweichungen darzulegen.

¹⁸² DANIEL KAHNEMAN, New Challenges to the Rationality Assumption, Journal of institutional and theoretical economics 1994, 18-36, 25 f.

§ 4 Rezeption der psychologischen Forschung durch die Rechtswissenschaft:

Behavioral Law and Economics oder Empirical Legal Realism

- 84 MICHAEL J. SAKS und ROBERT F. KIDD waren 1980 die ersten Juristen, die sich explizit auf die sozialpsychologische Forschung zu den *heuristics and biases*, Rückschaffehler und Overconfidence beriefen.¹⁸³ Sicherlich beeinflusst vom sehr pessimistischen Bild menschlichen Denkens, das in den siebziger Jahren gezeichnet worden war, empfahlen sie, bei der Feststellung von Tatsachen durch Gerichte mathematische Hilfsmittel (*decision aids*) bei zu ziehen (der Artikel war eine durch die psychologische Forschung untermauerte Antwort auf LAWRENCE TRIBES berühmten Aufsatz „*Trial by Mathematics: Precision and Ritual in the Legal Process*“, in dem TRIBE 1971 die Anwendung des Bayes-Theorem in Geschworenenprozessen abgelehnt hatte).¹⁸⁴ CATHERINE FITZMAURICES und KEN PEASES Buch „*The Psychology of Judicial Sentencing*“ aus dem Jahr 1986 enthielt ein Kapitel über „*Cognitive Errors and Judges*“, in dem sich die Autoren Gedanken über den Einfluss verschiedener kognitiver Täuschungen auf richterliche Entscheidungen machten.¹⁸⁵ Den Arbeiten von SAKS/KIDD und FITZMAURICE/PEASE gemeinsam ist, dass sie sich nicht auf die Ökonomische Analyse des Rechts beziehen und keine eigene empirische Forschung vorstellen.
- 85 Ein weiterer wichtiger früher Aufsatz, der ebenfalls 1986 erschien, stammt von WARD EDWARDS und DETLOF VON WINTERFELDT.¹⁸⁶ Die Autoren fassten die enorm umfangreiche sozialpsychologische Literatur zusammen, machten sie für Juristen zugänglich und wiesen auf die Relevanz dieser Forschungsergebnisse für das Rechtssystem hin (der Titel des Aufsatzes lautete „*Cognitive Illusions and their Implications for the Law*“). Da einer der Autoren zu den „Vätern“ der Forschung zu den kognitiven Illusionen gehörte, wurde der Aufsatz wahrgenommen. Auch EDWARDS und WINTERFELDT bezogen sich nicht explizit auf das der ÖAR zu Grunde liegende Menschenbild.

I. Behavioral Law and Economics

- 86 1989 machte ELLICKSON erstmals ausdrücklich die Verbindung zur Ökonomischen Analyse des Rechts. In einem viel beachteten Aufsatz plädierte er dafür, das Menschenbild der ÖAR durch menschliche Fehler und menschliche Kultur zu bereichern.¹⁸⁷ Interessant ist seine Argumentation, die direkt an das rationale Eigeninteresse junger Akademiker appellierte: er wies darauf hin, dass die analytischen Modelle der ÖAR inzwischen eine Komplexität erreicht hatten, die es einem Juristen ohne vertiefte mathematische Ausbildung kaum mehr erlaubt, einen echten Beitrag zur Forschung zu leisten. Gleichzeitig war

¹⁸³ MICHAEL J. SAKS/ROBERT F. KIDD, Human Information Processing and Adjudication: Trial by Heuristics, *Law & Society Review* 1980, 123-160.

¹⁸⁴ LAWRENCE TRIBE, *Trial by Mathematics: Precision and Ritual in the Legal Process*, *Harvard Law Review* 1971, 1329-1351.

¹⁸⁵ CATHERINE FITZMAURICE/KEN PEASE, *The Psychology of Judicial Sentencing*, Manchester 1986, 17-34.

¹⁸⁶ EDWARDS/WINTERFELDT, FN 178.

¹⁸⁷ ELLICKSON, FN 82.

der Einfluss der kognitiven Täuschungen noch weitgehend unerforscht, und die daraus zu gewinnenden Erkenntnisse höchstwahrscheinlich grösser als diejenigen, die durch eine weitere Verfeinerung der mathematischen Modelle erzielt werden konnten; umso eher als diese aus Sicht vieler Juristen immer weniger mit der Realität zu tun hatten. Der Grenznutzen des Studiums der sozialpsychologischen Literatur, so ELLICKSON, sei daher für junge Juristen, die eine akademische Karriere anstreben, grösser als derjenige eines Wirtschaftsstudiums, das notwendig wäre, um in der ÖAR wirklich mitreden zu können.¹⁸⁸

- 87 Sei es wegen des Aufrufs von ELLICKSON, sei es, weil die traditionelle Ökonomische Analyse des Rechts zunehmend an ihre Grenzen stiess, auf jeden Fall stieg in den neunziger Jahren des letzten Jahrhunderts die Zahl der Arbeiten, die versuchten, die Erkenntnisse der Forschung zu den kognitiven Täuschungen auf spezifisch juristische Fragestellungen anzuwenden, sprunghaft an.¹⁸⁹ Haftpflicht-,¹⁹⁰ Familien-,¹⁹¹ Vertrags-,¹⁹² Straf-,¹⁹³ Steuer-,¹⁹⁴ Börsen-,¹⁹⁵ Umweltrecht¹⁹⁶ u. v. a. wurden aus dem Blickwinkel der kognitiven Täuschungen analysiert.
- 88 Bis 1998 war die Literatur in einem Ausmass angeschwollen, die es angezeigt erschienen liess, die Gemeinsamkeiten der bisher erschienenen Arbeiten herauszuarbeiten. JOLLS, SUNSTEIN und THALER unternahmen diesen Versuch mit „*A Behavioral Approach to Law and Economics*“.¹⁹⁷ Kurz darauf brachten KOROBKIN und ULEN das Hauptanliegen der Strömung bereits im Titel ihrer Arbeit auf den Punkt: „*Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics*“.¹⁹⁸ Im gleichen Jahr veröffentlichte CASS R. SUNSTEIN einen Sammelband, der die wichtigsten Arbeiten der letzten zehn Jahre wiedergab und den Titel „*Behavioral Law & Economics*“ trug. Unter diesem Namen sollte die „Bewegung“ in Zukunft bekannt sein.¹⁹⁹ Kritiker haben bemerkt,

¹⁸⁸ ELLICKSON, FN 82, 33.

¹⁸⁹ Einen ausgezeichneten Überblick bietet DONALD C. LANGEVOORT, Behavioral Theories of Judgement and Decision Making in Legal Scholarship: A Literature Review, Vanderbilt Law Review 1998, 1499-1540.

¹⁹⁰ JEFFREY RACHLINSKI, A Positive Psychological Theory of Judging in Hindsight, University of Chicago Law Review 1998, 571-625.

¹⁹¹ LYNN A. BAKER/ROBERT E. EMERY, When Every Relationship is Above Average: Perceptions and Expectations of Divorce at the Time of Marriage, Law and Human Behavior 1993, 439-450.

¹⁹² MELVIN A. EISENBERG, The Limits of Cognition and the Limits of Contract, Stanford Law Review 1995, 211-259.

¹⁹³ NEAL KUMAR KATYAL, Deterrence's Difficulty, Michigan Law Review 1997, 2385-2476; LEE ROSS/DONNA SHESTOWSKY, Contemporary Psychology's Challenges to Legal Theory and Legal Practice, Northwestern University Law Review 2003, 1081-1114.

¹⁹⁴ EDWARD J. MCCAFFREY, Cognitive Theory and Tax, UCLA Law Review 1994, 1961-1981; Nachdruck in: SUNSTEIN (Hrsg.), 398-421.

¹⁹⁵ DONALD C. LANGEVOORT, Organized Illusions: A Behavioral Theory of Why Corporations Mislead Stock Market Investors (and Cause Other Social Harms), University of Pennsylvania Law Review 1997, 101-172; Nachdruck in: SUNSTEIN (Hrsg.), 144-167.

¹⁹⁶ ROGER C. NOLL/JAMES E. KRIER, Some Implications of Cognitive Psychology for Risk Regulation, Journal of Legal Studies 1990, 747-779.

¹⁹⁷ JOLLS/SUNSTEIN/THALER, FN 16.

¹⁹⁸ KOROBKIN/ULEN, FN 104.

¹⁹⁹ Bereits 1997 hatte SUNSTEIN einen Artikel mit dem Titel „Behavioral Analysis of Law“ veröffentlicht (University of Chicago Law Review 1997, 1175-1195). M. E. trifft der Titel des Aufsatzes den Kern der neuen Strömung besser als der Titel des Buches, da er sich nicht direkt auf die Ökonomik bezieht.

dass die Bezeichnung nicht passend sei: erstens sei „*behavioral*“ irreführend, da auf die psychologische Schule von WATSON/SKINNER verweisend, zweitens könne kaum mehr von „*economics*“ gesprochen werden, wenn die Rationalitätsannahme aufgegeben werde und drittens befasse sich die Strömung in erster Linie mit Entscheidungen.²⁰⁰ MITCHELL schlägt daher vor, statt dessen von „juristischer Entscheidungstheorie“ (*legal decision theory*) zu sprechen, da diese Bezeichnung klar mache, dass sich um die Anwendung der (psychologischen) Entscheidungstheorie auf rechtliche Sachverhalte handle und die meisten Arbeiten zudem ohnehin sehr theoretisch seien (eine kleine Spitze).²⁰¹ Da sich die Bezeichnung *Behavioral Law and Economics* eingebürgert hat, werde ich sie hier verwenden.

- 89 KOROBKINS und ULENS Hauptaussage ist, dass „dünne“ Modelle rationalen Verhaltens, mit wenig restriktiven Prämissen, ungeeignet sind zur Analyse rechtlicher Phänomene, da sie *ex ante* keine Voraussagen über das von den Rechtsunterworfenen zu erwartende Verhalten erlauben. Andererseits seien die Voraussagen der „dichten“ Modelle mit stringenten Annahmen, oft schlicht falsch.²⁰² Um das eine Problem – mangelnde Prognosefähigkeit – zu beheben, ohne das andere Problem – unrichtige Prognosen – zu verschärfen, müsse das der ÖAR zu Grunde liegende Modell daher um die Erkenntnisse der sozialpsychologischen Forschung ergänzt werden und die ÖAR vermehrt empirisch forschen.²⁰³
- 90 JOLLS et al. bringen eine sehr ähnliche Kritik am traditionellen Menschenbild der ÖAR vor wie KOROBKIN/ULEN. Sie weisen darauf hin, dass es die Grenzen menschlicher Kognition fraglich erscheinen lassen, dass die oft anti-interventionistischen Ratschläge von Vertretern der ÖAR optimal sind, da Menschen in Anbetracht ihrer Grenzen manchmal vor sich selber geschützt werden müssen.²⁰⁴ Sowohl JOLLS et al. als auch KOROBKIN/ULEN räumen der Antwort auf die (antizipierten) Kritik an der BLE, auf die ich auf S. 47 ff. näher eingehen werde, viel Raum ein.
- 91 Als wissenschaftliche „Revolution“ oder „Paradigmawechsel“ kann man die BLE nicht bezeichnen. Es handelt sich um eine evolutionäre Weiterentwicklung der traditionellen Ökonomischen Analyse des Rechts, die zunehmend an ihre Grenzen stösst resp. für Juristen kaum mehr zugänglich ist. Die BLE übernimmt Pfeiler des ökonomischen Paradigmas, wie den methodologischen Individualismus und die Effizienz als Rechtsziel, während sie das Modell des Menschen modifiziert. Da sie von der Fehlbarkeit des Menschen ausgeht, steht sie „paternalistischen“ Lösungen tendenziell aufgeschlossener gegenüber als Vertreter der traditionellen ÖAR, die das Effizienzziel oft am besten verwirklicht sehen, wenn den Marktkräften freien Lauf gelassen wird.²⁰⁵ Da Menschen erwiesenermassen bei freier Wahl oft nicht die Wahl treffen, die in ihrem besten Interesse liegt, ist es für Vertreter der BLE

²⁰⁰ GREGORY A. MITCHELL, Why Law and Economics' Perfect Rationality Should Not Be Traded for Behavioral Law and Economics' Equal Incompetence, Georgetown Law Journal 2002, 67-167, 78 f.

²⁰¹ MITCHELL, FN 200, 79.

²⁰² KOROBKIN/ULEN, FN 104, 1069 f.

²⁰³ KOROBKIN/ULEN, FN 104, 1058.

²⁰⁴ JOLLS/SUNSTEIN/THALER, FN 16, 46 f.

²⁰⁵ CASS R. SUNSTEIN/RICHARD H. THALER, Libertarian Paternalism Is Not an Oxymoron, University of Chicago Law Review 2003, 1159-1202.

gerechtfertigt, sie auf die richtige Wahl „hinzusteuern“, ohne dass geradezu Zwang ausgeübt wird oder gewisse Optionen gänzlich verboten werden.²⁰⁶

- 92 Natürlich ist es noch zu früh, eine Beurteilung des Erfolgs von *Behavioral Law and Economics* abzugeben. Wie POSNER nicht ganz zu Unrecht bemerkt, hat die Ökonomische Analyse des Rechts schon andere Attacken überstanden.²⁰⁷ Klar ist, dass sich die (amerikanischen) Juristen in den letzten zehn Jahren immer stärker mit den Forschungen von KAHNEMAN und TVERSKY auseinandergesetzt haben. Abbildung 5 zeigt das Verhältnis der Treffer von [kahneman & tversky] zu „*economic analysis of law*“ in der *LexisNexis* Datenbank *Law Journals* seit 1980. Während 1980-84 auf zehn Erwähnungen von „*economic analysis of law*“ eine Erwähnung von „KAHNEMAN & TVERSKY“ gekommen ist, ist dieses Verhältnis bis zum Jahr 2003 fast auf 2:1 gestiegen.

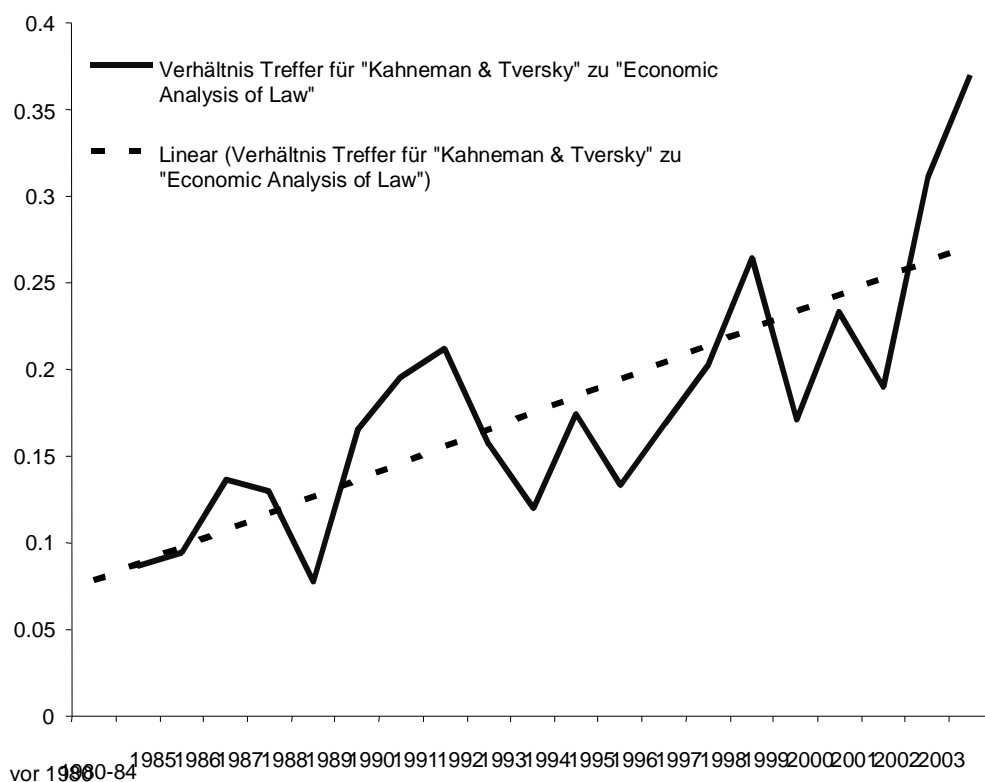


Abbildung 5: Verhältnis Treffer „Kahneman & Tversky“ zu „Economic Analysis of Law“

²⁰⁶ SUNSTEIN/THALER, FN 205, 1163; COLIN CAMERER/SAMUEL ISSACHAROFF/GEORGE LOEWENSTEIN/TED O'DONOGHUE/MATTHEW RABIN, Regulation for Conservatives: Behavioral Economics and the Case for "Asymmetric Paternalism", *University of Pennsylvania Law Review* 2003, 1211-1254, 1212. Kritisch JEFFREY J. RACHLINSKI, The Uncertain Psychological Case for Paternalism, *Northwestern University Law Review* 1997, 1165-1225.

²⁰⁷ POSNER, FN 67, 28.

- 93 Als weiteres quantitatives Indiz für den Erfolg der neuen Bewegung kann man ansehen, dass einer ihrer prominentesten Vertreter der meistzitierte (amerikanische) Jurist unter 50 Jahren ist (Tabelle 3).

<i>Name</i>	<i>Zitate</i>
Cass R. Sunstein (1954)	2,701
Frank H. Easterbrook (1948)	2,432
Daniel A. Farber (1950)	1,536
A. Mitchell Polinsky (1948)	1,245
Daniel R. Fischel (1950)	1,095

Tabelle 3: Meistzitierte Juristen unter 50 Jahren (aus SHAPIRO, 2000)

II. Empirical Legal Realism?

- 94 DANIEL A. FARBER hat die neue Strömung, die versucht, sozialpsychologische Forschung vermehrt für die Analyse juristischer Probleme fruchtbar zu machen, als erster explizit in die Tradition des *Legal Realism* gestellt, als er seine Besprechung von SUNSTEINS Aufsatzsammlung mit „*Towards A New Legal Realism*“ betitelte.²⁰⁸ Er kritisiert an den *Behavioral Law and Economists*, dass sie sich wie die Vertreter der traditionellen ÖAR zu sehr auf die Analyse formalisierter Wahlprobleme zwischen meist genau definierten Alternativen beschränken. In Wirklichkeit seien viele Entscheidungssituationen nicht klar strukturiert, und wie der Entscheidende die Entscheidungssituation konstruiere, sei häufig wichtiger als die Wahl selbst. Zweitens neigten die *Behavioral Economists* wie die orthodoxen Ökonomen dazu, ihr Modell menschlichen Verhaltens undifferenziert auf zahlreiche unterschiedliche Lebenssachverhalte anzuwenden und dabei den kulturellen und sozialen Kontext ausser Acht zu lassen. Drittens seien sie sich der normativen Herausforderungen, die ihr neues Modell menschlichen Verhaltens mit sich bringe, wenig bewusst und verliessen sich zu unkritisch auf die normativen Kriterien der Wohlfahrtsökonomik, die nicht unabhängig vom *rational choice* Modell angewandt werden könnten.²⁰⁹
- 95 Für FARBER muss die Bewegung daher ihren Horizont erweitern und weitere sozialwissenschaftliche Erkenntnisse berücksichtigen, den institutionellen Kontext der Entscheidungen verstärkt beachten und die normativen Herausforderungen, die sich aus der Beobachtung ergeben, dass Präferenzen nicht stabil sind, durchdenken.²¹⁰ SUNSTEIN und Kollegen könnten dann als Wegbereiter eines wieder belebten *Legal Realism*, der sich auf empirische Forschung stützen kann, betrachtet werden.²¹¹ In eine ähnliche Kerbe hatte bereits KELMAN geschlagen, als er von einem „rhetorischen Duett“ von ÖAR und BLE sprach, wenn es den Vertretern der BLE nicht gelänge, sich aus dem Korsett der Analyse genau definierter

²⁰⁸ DANIEL A. FARBER, *Towards A New Legal Realism*, University of Chicago Law Review 2001, 279-303.

²⁰⁹ FARBER, FN 208, 296.

²¹⁰ FARBER, FN 208, 302.

²¹¹ FARBER, FN 208, 303.

Wahlsituationen zu lösen.²¹² Ansonsten bliebe *Behavioral Law and Economics* eine grösstenteils parasitäre, die Ökonomische Analyse des Rechts kritisierende Strömung.²¹³

- 96 Im Sommer 2003 fand an der *Northwestern University* ein Symposium unter dem Titel „*Empirical Legal Realism: A New Social Scientific Assessment of Law And Human Behavior*“ statt, an dem führende Vertreter der BLE wie GUTHRIE, KOROBKIN, RACHLINSKI und SUNSTEIN teilnahmen.²¹⁴ Sozialwissenschaftliche Erkenntnisse sollen, wie ursprünglich die Rechtsrealisten gefordert hatten, zum Verständnis des Rechts bei gezogen werden.²¹⁵ „Empirisch“ ist dieser Rechtsrealismus insofern, als er sich auf empirische psychologische Forschung stützen kann und, wie die Arbeiten von GUTHRIE et al.²¹⁶, KELMAN et al.²¹⁷, KOROBKIN²¹⁸, RACHLINSKI²¹⁹, SUNSTEIN et al.²²⁰ und anderen zeigen, Juristen vermehrt selber empirisch forschen. Die Abwendung von der Bezeichnung *Behavioral Law and Economics* kann als Zeichen gedeutet werden, dass man sich vermehrt von der Ökonomischen Analyse des Rechts abgrenzen und als eigenständige Disziplin wahrgenommen werden will. Die Bezeichnung *Empirical Legal Realism* bietet auch die Chance, sich von der reinen Kritik an der rationalen Entscheidungstheorie zu lösen und die Bewegung breiter abzustützen.

²¹² MARK KELMAN, Behavioral Economics as Part of a Rhetorical Duet: A Response to Jolls, Sunstein and Thaler, *Stanford Law Review* 1998, 1577-1591, 1591.

²¹³ KELMAN, FN 212, 1586.

²¹⁴ Die Beiträge sind in der *Northwestern University Law Review* 2003, 1075 ff., publiziert; eine kurze Zusammenfassung jedes Beitrages enthält ROSS/SHESTOWSKY, FN 193.

²¹⁵ FARBER, FN 208, 303.

²¹⁶ GUTHRIE/RACHLINSKI/WISTRICH, FN 20.

²¹⁷ MARK KELMAN/YUVAL ROTTENSTREICH/AMOS TVERSKY, Context-Dependence in Legal Decision Making, *Journal of Legal Studies* 1996, 287-318; Nachdruck in: SUNSTEIN (Hrsg.), 61-94.

²¹⁸ RUSSEL KOROBKIN/CHRIS GUTHRIE, Opening Offers and Out-of-Court Settlement: A Little Moderation May not Go a Long Way, *Ohio State Journal on Dispute Resolution* 1994, 1-20.

²¹⁹ JEFFREY J. RACHLINSKI, Gains, Losses and the Psychology of Litigation, *Southern California Law Review* 1996, 113-185.

²²⁰ CASS R. SUNSTEIN/DANIEL KAHNEMAN/DAVID SCHADKE, Assessing Punitive Damages (With Notes on Cognition and Valuation in Law), *Yale Law Journal* 1998, 2071-2152; Nachdruck in: SUNSTEIN (Hrsg.), 232-258.

§ 5 Kritik

I. Kritik an der psychologischen Forschung zu den *heuristics and biases*

- 97 Der *heuristics and biases* Ansatz wird kritisiert, weil er sich einseitig auf die Fehler im Sinne einer Abweichung von einer (allzu) eng verstandenen Rationalität konzentriere, statt die adaptiven Vorteile der Urteilsheuristiken zu erforschen,²²¹ weil er den Kontext von Entscheidungen vernachlässige und daher Laborresultate leichtfertig generalisiere und weil er in 30 Jahren Forschung nicht geschafft habe, eine Theorie zu formulieren, die sowohl fehlerhafte wie korrekte Entscheidungen voraussagen kann.²²²

A. Zu pessimistische Sicht

- 98 CHRISTENSEN-SZALANKSI und BEACH stellten 1984 fest, dass Studien, die sich mit menschlichen Urteilen, Entscheidungen und Problemlösungsverhalten befassten, signifikant häufiger zitiert werden, wenn sie einen „Fehler“ menschlichen Denkens finden, als wenn sie zum Schluss kommen, dass die Versuchspersonen die Aufgabe im grossen und ganzen gut gelöst hatten.²²³ Dieser Schluss wird von ROBINS und CRAIK allerdings in Zweifel gezogen; nach ROBINS/CRAIK ist der beobachtete Unterschied bei sorgfältiger statistischer Auswertung gering und lässt sich dadurch erklären, dass die pessimistischen Studien in anerkannteren Zeitschriften (d.h. solchen mit höherem „*impact factor*“) publiziert wurden.²²⁴ Klar ist, dass Fehler menschlicher Entscheidungsstrategien, gerade für ein Publikum ausserhalb der Psychologie, interessanter sind als die Feststellung, dass Menschen meist richtig entscheiden. „*If it bleeds, it leads*“ gilt nicht nur im Journalismus, sondern auch in der wissenschaftlichen Textproduktion und -rezeption.²²⁵ Die Kritik an der pessimistischen Sicht des *heuristics and biases* Paradigmas wurde immer wieder vorgebracht,²²⁶ aber auf Grund der „Nachfrage“ nach entsprechenden Resultaten wird sich an dieser Sicht so bald wahrscheinlich nichts ändern. TVERSKY und KAHNEMAN haben zwar immer wieder betont, dass das *heuristics and biases* Programm eine negative wie eine positive Agenda habe; negativ in dem Sinn, dass es nach Fehlern sucht, positiv, weil es die aus dem Studium der Fehler gewonnen Erkenntnisse verwendet, um die ihnen zu Grunde liegenden Denkprozesse zu verstehen.²²⁷ Allerdings kann das nicht darüber hinwegtäuschen, dass das Schwergewicht auf der Beschreibung und Katalogisierung der Fehler lag.

²²¹ LOLA L. LOPES, *The Rhetoric of Irrationality*, *Theory & Psychology* 1991, 65-82.

²²² GERD GIGERENZER, *How to Make Cognitive Illusions Disappear: Beyond "Heuristics and Biases"*, *European Review of Social Psychology* 1991, 83-115, 101.

²²³ JAY J. CHRISTENSEN-SZALANSKI/LEE ROY BEACH, *The Citation Bias: Fad and Fashion in the Judgment and Decision Literature*, *American Psychologist* 1984, 75-78.

²²⁴ RICHARD W. ROBINS/KENNETH CRAIK, *Is There A Citation Bias in the Judgement and Decision Literature?*, *Organizational Behavior and Human Decision Processes* 1993, 225-244.

²²⁵ GILOVICH/GRIFFIN, FN 151, 8.

²²⁶ Zuletzt von JOACHIM L. KRUEGER/DAVID C. FUNDER, *Towards a Balanced Social Psychology: Causes, Consequences, and Cures for the Problem Seeking Approach to Social Behavior and Cognition*, *Behavioral and Brain Sciences* 2004, 313-376.

²²⁷ Nachweise bei GILOVICH/GRIFFIN, FN 151, 4.

B. Der „richtige“ normative Standard

- 99 Die Diskussion um den richtigen normativen Standard konzentriert sich vor allem auf *bias*es im Zusammenhang mit statistischem Denken, wie der Vernachlässigung der Basisrate, Verfügbarkeitsheuristik, dem Konjunktions-Fehler (*conjunction fallacy*) und Overconfidence.²²⁸ Umstritten ist, ob (nur) aus der Anwendung des Bayes-Theorem die korrekte Antwort auf die gestellten Fragen resultiert,²²⁹ ob Vernachlässigung der Basisrate und Overconfidence verschwinden, wenn die zu verarbeitende Information in der Form natürlicher Häufigkeiten dargestellt wird²³⁰ oder ob die Versuchspersonen eine richtige Antwort auf missverständlich gestellte Fragen geben.²³¹ Die Kritik scheint teilweise gerechtfertigt, insbesondere, was die Repräsentation der Information angeht, betrifft aber nur zwei der Heuristiken (Repräsentativität) und Fehler (Overconfidence), die in dieser Arbeit besprochen werden. Ich werde bei der Darstellung dieser beiden Phänomene näher auf die Kritik eingehen; insbesondere die Frage, ob das Bayes-Theorem ein angemessener normativer Standard ist, kann nicht auf wenigen Zeilen abgehandelt werden.

C. Generalisierbarkeit

- 100 Welche der *heuristics and biases* sich auf Entscheidungen in der „wirklichen“ Welt generalisieren lassen, ist eine Frage, die noch ungenügend erforscht ist, wie auch Anhänger des *heuristics and biases* Programms zugeben.²³² Sie weisen immerhin darauf hin, dass KAHNEMAN und TVERSKY nicht Laborexperimente konstruiert haben, um Fehler ihrer Versuchspersonen nachweisen zu können, sondern weil sie beobachteten, dass ihre Studenten bestimmte statistische Gesetze nicht oder nur sehr schwer verstanden und sie versuchten, diesen Fehlern mit ihren Experimenten auf die Schliche zu kommen.²³³ M. a. W. stand am Anfang des Forschungsprogramms die Beobachtung, dass ausserhalb des Labors gewisse Fehler systematisch und immer wieder gemacht wurden. Erst dann wurde die Forschung ins Labor verlegt. Dass anfänglich kaum Versuche unternommen wurden, die im Labor gewonnen Resultate ausserhalb des Labors zu validieren, liegt am Aufwand und den Kosten, den solche Studien mit sich bringen, und dem geringen Prestige, der mit ihnen verbunden ist. Akademische Karriere macht, wer neue Fehler entdeckt, alte Forschungsergebnisse widerlegt oder neue Theorien entwirft, die die Fehler – und hoffentlich neue, noch unerforschte Phänomene – voraussagen.²³⁴ Erst wenn sich ein Phänomen im Labor unter verschiedenen Bedingungen als robust erwiesen hat, lohnt es sich, zu erforschen, ob es auch ausserhalb des Labors robust ist. Inzwischen wurden doch einige Feldstudien mit

²²⁸ GIGERENZER, FN 222.

²²⁹ Kritisch GERD GIGERENZER, Die Repräsentation von Information und ihre Auswirkung auf statistisches Denken, in: HELL/FIEDLER/GIGERENZER (Hrsg.), FN 11, 99-127, 117 ff.

²³⁰ So GIGERENZER, FN 222, 89 ff.; a.M. KAHNEMAN/TVERSKY, FN 179; GILOVICH/GRIFFIN, FN 151, 14.

²³¹ BARBARA MELLERS/RALPH HERTWIG/DANIEL KAHNEMAN, Do Frequency Representations Eliminate Conjunction Effects?, *Psychological Science* 2001, 269-275.

²³² GILOVICH/GRIFFIN, FN 151, 8.

²³³ GILOVICH/GRIFFIN, FN 151, 11.

²³⁴ JAMES FRIEDRICH, The “Bias” Bias in Social Psychology: Adaptive When and How?, *Behavioral and Brain Sciences* 2004, 335-336, 336.

positiven Resultaten durchgeführt,²³⁵ so dass der Vorwurf der fehlenden externen Validität nicht mehr generell aufrechterhalten werden kann. Andere Autoren weisen nicht ganz zu Unrecht darauf hin, dass sich die natürliche Umwelt des Menschen immer mehr Laborbedingungen angleicht; d.h. die Fähigkeit, abstrakte „Denksportaufgaben“ zu lösen, wie sie in typischen Laborexperimenten gestellt werden, mag im Pleistozän wenig wichtig gewesen sein, ist aber in einer modernen, komplexen, technologisierten, sich schnell verändernden Gesellschaft zunehmend zentral.²³⁶

D. Theoriedefizit

- 101 Insbesondere GIGERENZER kritisiert das *Theoriedefizit* des *heuristics and biases* Ansatzes.²³⁷ In 30 Jahren sei man nicht über die Beschreibung der Phänomene hinausgekommen, die zudem oft so vage definiert seien, dass mit ihnen alles erklärt werden könne.²³⁸ Was not tue, sei die Entwicklung von exakten und falsifizierbaren Modellen.²³⁹ TVERSKY und KAHNEMAN meinen, dass ein Begriff wie „Repräsentativität“ nicht a priori definiert werden muss, da experimentell nachgewiesen werden kann, wie repräsentativ eine Instanz in den Augen der Versuchspersonen für die Kategorie ist, der sie angehört.²⁴⁰ In der Modellierung liege nicht der Königsweg für den Fortschritt der Psychologie, die qualitativen Gesetze der Gestaltpsychologie²⁴¹ hätten verfrühte Versuche der Modellierung auch überlebt.²⁴²
- 102 In dem Vorwurf der Vagheit der Begriffe und der insgesamt fehlenden Exaktheit des *heuristics and biases* Ansatzes liegt sicher ein Körnchen Wahrheit. Für den psychologischen Laien liegt darin aber natürlich auch ein Teil der Attraktivität des Gebietes, und die fehlende Modellierung dürfte zum Erfolg der Bewegung ausserhalb der Psychologie beitragen haben. Psychologische Modelle haben zudem oft den Nachteil, dass sie entweder so abstrakt sind, dass sie kaum spezifische Voraussagen erlauben, oder dann so spezifisch sind, dass sie nur gerade ein bestimmtes Experiment modellieren.²⁴³ Die Forderung nach der Erforschung der Denkprozesse, die den beobachteten Fehlern zu Grunde liegen, ist aus Sicht der Psychologie verständlich. Tatsächlich erscheint es wenig befriedigend, nach 30 Jahren immer noch weitere Ausstellungsstücke ins ohnehin bereits überfüllte „Museum der Inkompetenz“²⁴⁴ zu stellen. Für den Juristen oder Ökonomen aber ist es möglicherweise

²³⁵ Nachweise S. 75 ff. sowie bei den einzelnen kognitiven Täuschungen.

²³⁶ KEITH E. STANOVICH/RICHARD F. WEST, Individual Differences in Reasoning: Implications for the Rationality Debate, Behavioral and Brain Sciences 2000, 645-726, 663.

²³⁷ GIGERENZER, FN 222, 101 ff.; GERD GIGERENZER, On Narrow Norms and Vague Heuristics: A Rebuttal to Kahneman and Tversky, Psychological Review 1996, 592-596.

²³⁸ GIGERENZER, FN 237, 592.

²³⁹ GIGERENZER, FN 222, 104 f.

²⁴⁰ KAHNEMAN/TVERSKY, FN 230, 585.

²⁴¹ Zu diesen JOCHEN MÜSSELER, Visuelle Wahrnehmung, in: JOCHEN MÜSSELER/WOLFGANG PRINZ, Allgemeine Psychologie, Heidelberg/Berlin 2002, 38 f.

²⁴² KAHNEMAN/TVERSKY, FN 230, 591.

²⁴³ Das „*Realistic Accuracy Model*“ von DAVID C. FUNDER, On the Accuracy of Personality Judgment: a Realistic Approach, Psychological Review 1995, 652-670, dass er dem „bias“ Programm der Sozialpsychologie gegenüberstellt (KRUEGER/FUNDER, FN 226, 326), würde ich zur ersten Kategorie zählen.

²⁴⁴ KRUEGER/FUNDER, FN 226, 322.

weniger relevant, zu wissen, *weshalb* eine bestimmte kognitive Täuschung auftritt, als wie und unter welchen Bedingungen. Das Theoriedefizit des *heuristics and biases* Programms erscheint daher für einen Beobachter ausserhalb des Fachgebiets nicht als schwerer Mangel; ausser er hätte den – meines Erachtens zum Scheitern verurteilten Anspruch – das *rational actor* Modell durch ein ähnlich elegantes und komplettes deskriptives Modell menschlichen Verhaltens zu ersetzen.

E. Duales System

- 103 In jüngster Zeit hat KAHNEMAN postuliert, dass der Mensch zwei verschiedene Denkprozesse verwendet: einen regelbasierten, langsamen, bewussten Prozess und einen assoziativen, schnellen, automatischen Prozess.²⁴⁵ KAHNEMAN übernimmt die Terminologie von STANOVICH/WEST, die die regelbasierten Prozesse als „System 1“, die assoziativen als „System 2“ bezeichnen.²⁴⁶ Die Idee, zwischen Intuition und Vernunft zu unterscheiden, ist natürlich schon viel älter.²⁴⁷ KAHNEMAN ist zuversichtlich, dass mit diesem Modell widersprüchliche experimentelle Befunde erklärt werden können. Experimente, in denen kognitive Täuschungen verschwinden, erleichtern seiner Ansicht nach die Anwendung des regelbasierten Systems 2. Sie ermöglichen keine Aufschlüsse über die intuitiven Prozesse des Systems 1, welche nach wie vor ablaufen, aber vom regelbasierten System 2 korrigiert werden.²⁴⁸

System 1 (intuitiv)	System 2 (reflexiv)
<i>Eigenschaften der Denkprozesse</i>	
automatisch	beherrscht
müheless	anstrengend
assoziativ	deduktiv
schnell, parallel	langsam, seriell
unbewusst	bewusst
Anwendung einer Fähigkeit	Anwendung einer Regel
<i>Inhalte, die mit den Systemen verarbeitet werden</i>	
affektiv	neutral
Kausalzusammenhänge	Statistik
konkret, spezifisch	Abstrakt
Prototypen	Klassen (<i>sets</i>)

Tabelle 4: Eigenschaften der beiden kognitiven Systeme (nach KAHNEMAN/FREDERICK, 2002)

²⁴⁵ DANIEL KAHNEMAN/SHANE FREDERICK, Representativeness Revisited: Attribute Substitution in Intuitive Judgment, in: GILOVICH/GRIFFIN/KAHNEMAN (Hrsg.), 49-81, 51.

²⁴⁶ STANOVICH/WEST, FN 236, 658.

²⁴⁷ Nachweise bei STANOVICH/WEST, FN 236, 658; und STEVEN A. SLOMAN, Two Systems of Reasoning, in: GILOVICH/GRIFFIN/KAHNEMAN (Hrsg.), 379-396, 380 f.

²⁴⁸ DANIEL KAHNEMAN in seiner Autobiographie für das Nobelpreiskomitee (erhältlich unter <http://nobelprize.org/economics/laureates/2002/kahneman-autobio.html>; besucht am 14. April 2005).

- 104 Über das Verhältnis der beiden Systeme bestehen zwei grundsätzlich verschiedene Auffassungen. Nach der einen Auffassung verwenden Menschen absichtlich Prozesse des Systems 1, wenn die Entscheidung relativ unwichtig und die Motivation gering ist. Die aufwändigeren Prozesse des Systems 2 werden eingesetzt, wenn es um viel geht.²⁴⁹ Nach einer anderen Auffassung, die derjenigen von KAHNEMAN/FREDERICK entspricht und die herrschende sein dürfte, arbeiten die beiden Systeme parallel und gleichzeitig. Das assoziative, schnelle System 1 arbeitet immer, nicht nur, wenn die Motivation gering ist.²⁵⁰ System 2 überwacht die Ergebnisse des Systems 1 und akzeptiert oder korrigiert sie.²⁵¹
- 105 Eine normativ korrekte Entscheidung kann nach diesem Modell auf zwei Arten zustande kommen: die Ergebnisse des Systems 1 und 2 stimmen überein oder das System 2 korrigiert das (vom normativen Standard abweichende) Ergebnis des Systems 1. Inkorrekte Entscheidungen können ebenfalls auf zwei Arten entstehen: das System 1 kommt zu einem inkorrekten Ergebnis und wird entweder vom System 2 nicht korrigiert, oder das System 2 korrigiert ungenügend, so dass die Entscheidung eine Mischung der Ergebnisse der beiden Prozesse ist. Unter normalen Umständen akzeptiert das System 2 die Ergebnisse des Systems 1 dabei nach oberflächlicher Prüfung, wie folgendes Beispiel zeigt: „Ein Schokoladeriegel und ein Kaugummi kosten zusammen Fr. 1,10. Der Schokoladeriegel kostet Fr. 1 mehr als der Kaugummi. Wie viel kostet der Kaugummi?“. ²⁵² Die meisten Menschen antworten spontan „10 Rappen“, weil die Summe von Fr. 1,10 einleuchtend in Fr. 1 und 10 Rappen zerfällt und 10 Rappen von der Grössenordnung her ungefähr zu stimmen scheinen. Das System 2 korrigiert offenbar das System 1 nur ungenügend; die meisten Menschen, so KAHNEMAN/FREDERICK, „sind sich nicht gewohnt, fest zu denken und oft zufrieden mit einem plausiblen Urteil, das schnell in den Sinn kommt“. ²⁵³
- 106 Empirische Unterstützung erhält das Modell der zwei Systeme durch Experimente, die (wenig überraschend) zeigen, dass Versuchspersonen mehr falsche Antworten geben, wenn sie abgelenkt werden (*cognitive load*)²⁵⁴ oder unter Zeitdruck stehen.²⁵⁵ Das langsame und anstrengende System 2 ist in diesen Fällen offenbar nicht in der Lage, die Ergebnisse des Systems 1 zu korrigieren. Weitere Unterstützung erhält das Modell durch den Befund, dass Versuchspersonen, die die normativ richtige Antwort geben, in der Regel signifikant besser abschneiden bei Tests der analytischen Intelligenz.²⁵⁶ Intelligente Menschen profitieren von einem relativ effizienten System 2, das ihnen erlaubt, fehlerhafte Intuitionen zu erkennen und zu korrigieren. Statistisch gebildete Versuchspersonen schneiden ebenfalls besser ab als naive Versuchspersonen, allerdings nur, wenn das Experiment die statistische Natur der Frage transparent macht.²⁵⁷

²⁴⁹ Nachweise bei GILOVICH/GRIFFIN, FN 151, 16.

²⁵⁰ KAHNEMAN/FREDERICK, FN 245, 51.

²⁵¹ KAHNEMAN/FREDERICK, FN 245, 51.

²⁵² Leicht angepasst aus KAHNEMAN/FREDERICK, FN 245, 58.

²⁵³ KAHNEMAN/FREDERICK, FN 245, 58.

²⁵⁴ Beispielsweise müssen sich die Versuchspersonen an eine mehrstellige Zahl erinnern, die sie nicht aufschreiben dürfen.

²⁵⁵ Nachweise bei KAHNEMAN/FREDERICK, FN 245, 58.

²⁵⁶ STANOVICH/WEST, FN 236, 653 ff.

²⁵⁷ KAHNEMAN/FREDERICK, FN 245, 68.

- 107 Die Präsentation statistischer Informationen in der Form natürlicher Häufigkeit beeinflusst nach KAHNEMAN/FREDERICK nicht die Intuition des Systems 1, die nach wie vor von der Repräsentativität geleitet wird, sondern vereinfacht die Korrekturmechanismen des Systems 2.²⁵⁸ Allgemeiner formuliert verschwinden kognitive Täuschungen, je mehr die Versuchsanordnung die anzuwendende Regel transparent macht; auch dies ist nicht wirklich überraschend.
- 108 Das Modell der zwei Systeme kann sicherlich zahlreiche experimentelle Befunde erklären; allerdings geht dabei auch einiges von der prognostischen Fähigkeit des ursprünglichen *heuristics and biases* Programms verloren. *Ex post*, wenn das Experiment keine systematischen Fehler gezeigt hat, kann man dies natürlich immer der Korrektur durch das System 2 zuschreiben und postulieren, dass das System 1 aufgrund einer Urteilsheuristik entschieden hat, sich diese aber nicht in einem Fehler manifestiert, weil das Ergebnis durch das System 2 korrigiert wurde. Das Modell lässt sich kaum falsifizieren; zudem erläutern KAHNEMAN/FREDERICK nicht, wie man *ex ante* wissen kann, ob System 1 oder System 2 die Entscheidung bestimmen wird.
- 109 Das Modell der zwei Systeme wirft natürlich die Frage auf, inwiefern kognitive Täuschungen die Entscheidungen von Richtern und Parteien in Gerichtsverfahren beeinflussen. Sicherlich werden Entscheidungen durch Richter vom regelbasierten, langsamen, anstrengenden System 2 überprüft und korrigiert, so dass der Einfluss kognitiver Täuschungen ausgeschlossen werden kann? Sicherlich entscheiden Richter nicht intuitiv und schnell, oder?
- 110 Unter Juristen wie Laien ist die Vorstellung verbreitet, dass richterliche Urteile durch einen aufwändigen Prozess, nach langer Beratung durch bestens ausgebildete und erfahrene Entscheidungsträger, die zahlreiche Informationen berücksichtigen, gefällt werden.²⁵⁹ Richter betonen gerne, dass jeder Einzelfall komplex, einzigartig und unvergleichbar sei; eine Auffassung, die von Sozialpsychologen nicht geteilt wird.²⁶⁰ Selbst wenn die soziale Wirklichkeit unstreitig komplex ist, heisst das nicht, dass richterliche Entscheide komplex sind.²⁶¹ Dieser Eindruck der Komplexität entsteht wahrscheinlich durch Introspektion; niemand bezweifelt, dass Richter sich viele Gedanken machen über einen Fall und oft hin- und her gerissen sind. Jedoch lassen sich die Faktoren, die letztlich das Urteil beeinflussen, gemäss zahlreichen empirischen Studien an einer Hand abzählen.²⁶² Über die Strafzumessung wird meist schnell entschieden und wenig diskutiert,²⁶³ Entscheidungen, ob ein Angeeschuldigter auf Kautio n frei gelassen werden soll, werden auf Grund weniger Hinweise (*cues*) gefällt²⁶⁴ und die Höhe der Kautio n hängt praktisch nur vom Antrag des

²⁵⁸ KAHNEMAN/FREDERICK, FN 245, 69.

²⁵⁹ VLADIMIR J. KONECNI/EBBE B. EBBESEN, The Mythology of Legal Decision Making, International Journal of Law and Psychiatry 1984, 5-18, 5.

²⁶⁰ MARGRIT E. OSWALD, Psychologie richterlichen Strafens, Stuttgart 1994, 13.

²⁶¹ KONECNI/EBBESEN, FN 259, 6.

²⁶² KONECNI/EBBESEN, FN 259, 15.

²⁶³ FITZMAURICE/PEASE, FN 185, 2.

²⁶⁴ MANDEEP K. DHAMI/PETER AYTON, Bailing and Jailing the Fast and Frugal Way, Journal of Behavioral Decision Making 2001, 141-168, 161 f.

Staatsanwaltes ab.²⁶⁵ VAN DUYNE meint aufgrund eigener Beobachtungsstudien in den Niederlanden.²⁶⁶

Zusammengefasst lassen die Resultate der beiden Feldstudien an der Behauptung der Juristen zweifeln, Strafzumessung sei eine hochkomplexe Aufgabe, die sich mit einzigartigen Entscheidungen und einzigartigen Fällen befasst. Ohne behaupten zu wollen, die Aufgabe von Richtern und Strafverfolgern bestehe nur aus schnellen Schätzungen [*quick guesswork*] und Routine, dekoriert mit Prunk und Feierlichkeit, möchte ich sagen, dass juristische Entscheidungsfindung vergleichbar ist mit anderen Arten offen strukturierter Aufgaben, wie der Beurteilung von Kunstwerken durch eine Kommission oder der Benotung eines Aufsatzes durch einen Lehrer. Wenn Strafzumessung schwierig ist, dann wegen ihrer Ungewissheit, nicht wegen ihrer Komplexität.

- 111 Für KONECNI/EBBESEN ist die Komplexität und Einzigartigkeit juristischer Entscheidungen ein Mythos. Beide stellen wie VAN DUYNE fest, dass richterliche Gehirne den gleichen kapazitiven Beschränkungen unterliegen wie diejenigen der restlichen Bevölkerung, weshalb es nicht weiter überraschend sei, dass Richter eine begrenzte Anzahl von Informationen verarbeiten bei ihren Entscheidungen.²⁶⁷ Aus demselben Grund ist auch anzunehmen, dass Richter die gleichen Urteilsheuristiken verwenden wie die übrigen Menschen. Dass diese in jedem Fall durch das System 2 korrigiert werden, ist in Anbetracht des Zeitdruckes, unter dem Richter in der Regel stehen, nicht zu vermuten. Ausgenommen von dieser Annahme sind alle diejenigen Richter und Leser, die bei der Frage nach dem Preis des Kaugummis (S. 45) auf Anhieb die richtige Lösung gewusst haben.

II. Kritik an *Behavioral Law and Economics*

- 112 Die Diskussion um *Behavioral Law and Economics* in der juristischen Literatur folgt zu einem grossen Teil der Debatte zwischen Vertretern der traditionellen Ökonomik und *behavioral economists* ausserhalb des Rechts²⁶⁸ sowie der Debatte zwischen Anhängern der *heuristics and biases* Schule und ihren Gegnern in der Psychologie.²⁶⁹
- 113 Die Hauptargumente der Ökonomen lassen sich wie folgt zusammenfassen: da die Ökonomik an aggregiertem Verhalten interessiert ist, spielen individuelle Abweichungen vom Standard rationalen Verhaltens keine Rolle;²⁷⁰ der Wettbewerbsdruck wird irrationalen Verhalten ausmerzen;²⁷¹ ob ein Modell realistisch ist, ist irrelevant, solange es falsifizierbare Prognosen erlaubt, die meist zutreffen;²⁷² *behavioral economists* ersetzen

²⁶⁵ VLADIMIR J. KONECNI/EBBE B. EBBESEN, The Setting of Bail, in: VLADIMIR J. KONECNI/EBBE B. EBBESEN (Hrsg.), The Criminal Justice System: A Social-Psychological Analysis, San Francisco 1982, 191-230, 210 ff.

²⁶⁶ PETRUS VAN DUYNE, Simple Decision Making, in: DONALD C. PENNINGTON/SALLY LLOYD-BOSTOCK, The Psychology of Sentencing, Oxford 1987, 143-156, 151.

²⁶⁷ KONECNI/EBBESEN, FN 259, 15; VAN DUYNE, FN 266, 151 f.

²⁶⁸ RACHLINSKI, FN 17, 741.

²⁶⁹ MITCHELL, FN 200.

²⁷⁰ RICHARD A. POSNER, Rational Choice, Behavioral Economics, and the Law, Stanford Law Review 1998, 1551-1575, 1556 f.

²⁷¹ KIRCHGÄSSNER, FN 62, 211.

²⁷² FRIEDMAN, FN 130.

das kohärente System der traditionellen Ökonomik durch theoriefreie *ad hoc* Erklärungen.²⁷³ Die Hauptargumente der Psychologen lauten, wie bereits angetönt, dass das Bild, das in der *heuristics and biases* Literatur von den kognitiven Fähigkeiten des Menschen gezeichnet wird, viel zu pessimistisch ist und dass *behavioral law and economists* von einzelnen Laborexperimenten unzulässigerweise generalisieren, ohne dem Kontext der Entscheidung in der wirklichen Welt genügend Beachtung zu schenken.²⁷⁴

- 114 Das Argument, dass die ökonomische Theorie an aggregiertem Verhalten interessiert ist und daher individuelle Abweichungen vom rationalen Verhalten keine Rolle spielen, ist das schwächste Argument der Vertreter der traditionellen ÖAR. Es setzt voraus, dass die Abweichungen vom Standard zufällig verteilt sind und sich daher bei der Aggregation gegenseitig aufheben.²⁷⁵ Dies ist aber bei den kognitiven Täuschungen gerade nicht der Fall: es handelt sich um *systematische* und – zumindest unter Laborbedingungen – *voraus-sehbare* Abweichungen vom normativen Standard (Bayes-Theorem, Erwartungsnutzentheorie).²⁷⁶ Sie heben sich bei Aggregation nicht auf; die meisten Studien der *heuristics and biases* Literatur rapportieren aggregierte Ergebnisse, die vom normativen Standard abweichen – was von manchen Psychologen wiederum kritisiert wird, weil dadurch individuelle Unterschiede verborgen würden.²⁷⁷
- 115 Das Argument, dass der Markt irrationales Verhalten ausmerzen wird, hat für juristische Personen eine gewisse Plausibilität. Ein Unternehmen, das sich konstant irrational und daher per definitionem ineffizient verhält, wird über kurz oder lang von seinen Konkurrenten verdrängt werden.²⁷⁸ Bei natürlichen Personen ist das Argument jedoch wenig überzeugend. Wer sich „irrational“ im Sinne der ökonomischen Theorie verhält, stirbt nicht aus.²⁷⁹ Er lebt weiter und nimmt weiter am Marktgeschehen teil, so dass das Equilibrium auch von den Teilnehmern, die kognitiven Täuschungen unterliegen, beeinflusst wird. KOROBKIN/ULEN weisen zudem darauf hin, dass selbst angenommen, juristische Personen gehen Konkurs, wenn sie sich irrational verhalten, es doch plausibel ist, dass sie nicht sofort Konkurs gehen und sich zu einem beliebigen Zeitpunkt eine substantielle Anzahl der Marktteilnehmer nicht perfekt rational verhält.²⁸⁰ HANSON und KYSAR verwenden das Argument sogar gegen die traditionelle Ökonomik, in dem sie argumentieren, dass der Wettbewerb Firmen der Konsumgüterindustrie dazu treibt, die kognitiven Grenzen der Konsumenten bewusst oder unbewusst zum Vorteil der Firmen und zum

²⁷³ POSNER, FN 270, 1559 ff.

²⁷⁴ GREGORY MITCHELL, Taking Behavioralism Too Seriously? The Unwarranted Pessimism of the New Behavioral Analysis of Law, William & Mary Law Review 2002, 1907-2021.

²⁷⁵ KIRCHGÄSSNER, FN 62, 21.

²⁷⁶ ROBERT A. PRENTICE, Chicago Man, K-T Man, And the Future of Behavioral Law and Economics, Vanderbilt Law Review 2003, 1663-1777, 1725.

²⁷⁷ MITCHELL, FN 274, 1947 ff.

²⁷⁸ Etwas anderes ist es, einen Ruf zu haben, sich irrational zu verhalten, dies kann ein strategischer Vorteil sein; siehe AVINASH K. DIXIT/BARRY J. NALEBUFF, Thinking Strategically, New York/London 1993, 148 f.

²⁷⁹ KOROBKIN/ULEN, FN 104, 1070.

²⁸⁰ KOROBKIN/ULEN, FN 104, 1071.

Nachteil der Konsumenten auszunützen.²⁸¹ Kognitive Täuschungen spielten daher in der Praxis vielleicht eine noch wichtigere Rolle als im Labor, da bewusst Situationen kreiert würden, die leicht zu Täuschungen führen.

- 116 Die Behauptung, dass eine Theorie nicht an der Realität ihrer Prämissen, sondern an den Voraussagen, die sie generiert, gemessen werden soll – ohne dass man so weit gehen muss wie FRIEDMAN, der meint, dass die Realität der Annahmen überhaupt keine Rolle spiele²⁸² – wird auch von den Vertretern der BLE geteilt. Ihre Hauptkritik an der ÖAR ist nicht, wie noch diejenige von Kritikern in den siebziger Jahren, dass das Modell des *homo oeconomicus* „unrealistisch“ ist, sondern dass es erwiesenermassen zu falschen Prognosen führt. Die Laborexperimente von KAHNEMAN, TVERSKY und Co. vergleichen die Antworten der Versuchspersonen mit den Antworten, die gemäss *rational choice theory* zu erwarten wären, und kommen zum Schluss, dass sich die Versuchspersonen nicht wie vorhergesagt verhalten. Die Prämissen der traditionellen ÖAR werden also nicht kritisiert, weil sie unrealistisch sind, sondern weil sie nicht taugen, das Verhalten von echten Menschen zu modellieren.

- 117 Mehr dran ist am Vorwurf des Theoriedefizits der BLE. Am schärfsten hat diese Kritik wohl POSNER formuliert:²⁸³

Vertreter der Theorie des rationalen Handelns fragen sich, was der „*rational man*“ in einer bestimmten Situation machen würde, und gewöhnlich ist die Antwort ziemlich klar und kann mit dem tatsächlichen Verhalten verglichen werden, um zu prüfen, ob die Voraussage bestätigt wird. Manchmal wird die Voraussage nicht bestätigt – daher haben wir „*behavioral economics*“. Aber es ist zutiefst unklar, was der „*behavioral man*“ in einer bestimmten Situation machen würde. Er ist eine Mischung aus rationalen und nicht rationalen Fähigkeiten und Impulsen. Er könnte irgendetwas tun. JST [JOLLS, SUNSTEIN und THALER] haben weder eine kausale Erklärung für das Verhalten des „*behavioral man*“ noch ein Modell seiner Entscheidungsstrategie. [...] Die empirischen Misserfolge einer Theorie zu beschreiben, zu spezifizieren und zu klassifizieren ist eine wertvolle und wichtige wissenschaftliche Tätigkeit. Aber es ist keine alternative Theorie.

- 118 POSNER fährt fort mit der Bemerkung, dass mit der „Theorie“ der BLE alles erklärt werden könne.²⁸⁴

JSTs Theorie ist gefährlich nahe am Abgrund der Nicht-Falsifizierbarkeit, vielleicht ist sie schon hineingefallen. Wenn sich Menschen rational verhalten, behandeln JST das nicht als Widerspruch zur Annahme eingeschränkt rationalen Verhaltens. Wenn Menschen einer Versuchung widerstehen, wird das nicht als Widerspruch zur Annahme eingeschränkter Willenskraft angesehen. Und wenn sie sich egoistisch verhalten, wird das nicht als Widerspruch zur Annahme eingeschränkten Eigeninteresses angesehen. [...] Es taucht daher die Frage auf, welche Beobachtung JSTs Theorie falsifizieren könnte. Wenn es keine gibt, dann haben sie keine Theorie, sondern nur eine Sammlung von Herausforderungen für Theoretiker, in diesem Fall traditionelle [*rational choice*] Ökonomen [...].

²⁸¹ JON D. HANSON/DOUGLAS A. KYSAR, Taking Behavioralism Seriously: Some Evidence of Market Manipulation, Harvard Law Review 1999, 1420-1572; dies., Taking Behavioralism Seriously: The Problem of Market Manipulation, New York University Law Review 1999, 630-749.

²⁸² Zur Kritik daran PAUL A. SAMUELSON, Problems of Methodology – Discussion, American Economic Review Papers and Proceedings 1963, 231-236.

²⁸³ POSNER, FN 270, 1559.

²⁸⁴ POSNER, FN 270, 1560.

- 119 Die Vertreter der BLE streiten den Vorwurf, über keine ausformulierte Theorie menschlichen Verhaltens zu verfügen, nicht grundsätzlich ab. Zu ihrer Verteidigung bringen sie zwei Argumente vor. Bei BLE handle es sich um eine junge akademische Disziplin. Es sei normal, dass eine junge Disziplin zuerst die Fehler und Unzulänglichkeiten des herrschenden Paradigmas kritisiere, ehe sie in der Lage sei, der herrschenden Theorie eine eigene konkurrierende Theorie gegenüberzustellen. Das Ziel der verhaltenswissenschaftlichen Entscheidungstheorie sei auf jeden Fall nicht, die Misserfolge der traditionellen Ökonomik zu katalogisieren, sondern die Formulierung einer allgemeingültigen, deskriptiven Theorie menschlicher Entscheidungsstrategie.²⁸⁵ Die meines Erachtens überzeugendere Position ist diejenige der Skeptiker, die bezweifeln, dass es den Sozialpsychologen gelingen wird, eine Theorie zu entwickeln, die menschliches Verhalten allgemein besser beschreiben kann als die rationale Entscheidungstheorie. Das menschliche Verhalten ist aus dieser Sicht zu komplex, um durch eine einzige universelle Theorie (die gleichzeitig konkrete Voraussagen erlaubt) beschrieben zu werden, die so elegant, allgemeingültig und sparsam in den Annahmen ist wie das Modell des *homo oeconomicus*. Eher handle es sich bei der „Theorie“ der BLE um eine pragmatische Sammlung situationsspezifischer Verhaltensweisen.²⁸⁶ Die Strategie müsse darin bestehen, „Mini-Theorien“, oder Erklärungen und Voraussagen spezifischer juristischer Phänomene, zu entwickeln und empirisch zu überprüfen;²⁸⁷ was natürlich bedingt, dass Juristen vermehrt selber empirisch tätig werden.²⁸⁸ Der Nachteil, dass die Verhaltenswissenschaften keine einheitliche Theorie menschlichen Verhaltens zur Verfügung stellen könnten, sei im Recht weniger nachteilig als in der Wirtschaft, da die meisten Gesetze einen beschränkten Anwendungsbereich hätten und es daher genügen würde, wenn die Theorie in der Lage sei, das Verhalten der von diesem Gesetz spezifisch betroffenen Personen vorauszusagen.²⁸⁹ Was die BLE gegenüber der traditionellen ÖAR an Sparsamkeit (der Prämissen) und Allgemeinheit der Anwendbarkeit aufgebe, würde kompensiert durch den grösseren Nutzen, den die BLE bei der Analyse spezifischer juristischer Probleme bringe.²⁹⁰ BLE ist nach dieser Sicht eine pragmatische Wissenschaft, die auf ein grosses Theoriegebäude verzichtet und ihre Daseinsberechtigung nicht aus innerer Kohärenz und Allgemeingültigkeit, sondern aus äusserer Nützlichkeit bei der Analyse spezifischer Probleme herleitet.²⁹¹
- 120 Ökonomen haben die Ökonomik schon immer gerne mit der Physik verglichen.²⁹² „Physik“ steht dabei für mathematische Eleganz und präzise, quantitative Voraussagen.²⁹³ Mit der Ablösung der Physik als Leitwissenschaft durch die (Molekular-)Biologie steigt aber die Zahl der Ökonomen, die darauf hinweisen, dass ein Vergleich mit der Biologie vielleicht angebrachter wäre. HERBERT SIMON bemerkte, dass man von der Molekularbio-

²⁸⁵ RACHLINSKI, FN 17, 752.

²⁸⁶ KOROBKIN/ULEN, FN 104, 1075.

²⁸⁷ MITCHELL, FN 200, 132.

²⁸⁸ KOROBKIN/ULEN, FN 104, 1058.

²⁸⁹ KOROBKIN/ULEN, FN 104, 1072.

²⁹⁰ KOROBKIN/ULEN, FN 104, 1075.

²⁹¹ KOROBKIN/ULEN, FN 104, 1058.

²⁹² KIRCHGÄSSNER, FN 62, 258 ff.

²⁹³ Ein Bild der Physik, das sicherlich stark vom Schulunterricht geprägt ist. Moderne Physik ist eine primär empirische Wissenschaft.

logie lernen könne, dass nicht nur quantitative, sondern auch qualitative Strukturen und Erklärungen wissenschaftlichen Wert haben.²⁹⁴ Den Vertretern traditioneller ÖAR am Theoriedefizit der Verhaltenswissenschaftlichen Ökonomik und Recht antwortet FARBER mit folgendem Vergleich:²⁹⁵

Die letzte Verteidigungsstellung der Vertreter der rationalen Entscheidungstheorie ist, darauf zu bestehen, dass es eine Theorie brauche um eine Theorie zu schlagen; und dass die Verhaltenswissenschaftler nur eine Sammlung empirischer Unregelmässigkeiten ohne vereinigende Theorie hätten. Die Verhaltenswissenschaftler antworten entrüstet, sie hätten sehr wohl eine Theorie, zugegebenermassen eine unvollständige. Die Annahme ist offensichtlich auf beiden Seiten, die *conditio sine qua non* der Sozialwissenschaft sei eine einheitliche prädiktive Theorie. Aber vielleicht ist dies nur ein weiteres Symptom des berühmten „Physik-Neides“ der Ökonomik. Die Physik ist ein atemberaubendes Beispiel mathematischer Eleganz kombiniert mit fantastisch genauen Voraussagen. Aber die Physik als Paradigma der Wissenschaft zu betrachten, könnte ein Fehler sein. Die grösste Erfolgsgeschichte heute ist vielleicht die Biologie. Biologie hat ein zentrales Paradigma (Evolution) und ein Verständnis ihrer molekularen Basis. Aber Organismen sind [...] extrem komplex und niemand scheint zu glauben, dass ihre Eigenschaften im Detail auf der Basis einer deduktiven Theorie vorausgesagt werden können. Um ein zufälliges Beispiel aus einer aktuellen Zeitschrift zu nehmen: wer wissen will, welche Rolle Guanosin-Triphosphatase bei der Differenzierung der T-Zellen spielt, muss ins Labor gehen, um es herauszufinden. Niemand scheint zu glauben, dass man es im Voraus aufgrund einer allgemeinen Theorie berechnen kann. Manche finden die theoretische Eleganz der Physik sympathischer als die Komplexität der Molekularbiologie, und es wäre schön, wenn Sozialwissenschaft so elegant wäre wie die Relativitätstheorie oder Quantenphysik. Aber es gibt keinen Grund, anzunehmen, dass eine erfolgreiche Wissenschaft menschlichen Verhaltens mehr wie Physik als wie Biologie aussehen wird.

III. Stellungnahme

- 121 ARTHUR LEFF soll einmal gesagt haben, die Ökonomische Analyse des Rechts sei eine Wüste, Rechtssoziologie (*law and society*) aber ein Sumpf.²⁹⁶ BLE, die einen Platz zwischen ÖAR und Rechtssoziologie einnimmt,²⁹⁷ könnte die fruchtbare Savanne dazwischen sein. Es ist für eine junge Disziplin auf jeden Fall kein schlechtes Zeichen, wenn sie von rechts (ÖAR, POSNER) und links (Rechtssoziologie, KELMAN) kritisiert wird.
- 122 Mit der selbst gewählten Bezeichnung „*Behavioral Law and Economics*“ haben sich JOLLS, SUNSTEIN und THALER keinen Gefallen getan. Die Kritik an einem herrschenden Paradigma hat der Bewegung zwar viel Aufmerksamkeit und eine Struktur verschafft. Solange die Struktur aber darin besteht, Anomalien der rationalen Entscheidungstheorie aufzuzeigen, wird sie die Bewegung nicht während längerer Zeit tragen können. Die Hoffnung gewisser Vertreter der BLE, dass es der Psychologie in absehbarer Zeit gelingen wird, ein allgemeingültiges Modell menschlichen Verhaltens zu entwickeln, das der rationalen Entscheidungstheorie überlegen ist, teile ich nicht. Die Entwicklung wird eher in die von FARBER und MITCHELL aufgezeigte Richtung gehen, d.h. der Entwicklung situationsspezifischer Mini-Theorien.

²⁹⁴ SIMON, FN 143, 367.

²⁹⁵ FARBER, FN 208, 295 (interne Zitate weggelassen).

²⁹⁶ Zitiert nach ROBERT C. ELLICKSON, *Order Without Law*, Cambridge 1991, 147.

²⁹⁷ MARK KELMAN, *Law and Behavioral Science: Conceptual Overview*, Northwestern University Law Review 2003, 1347-1392, 1351.

- 123 Die Bewegung wird daher treffender als eine neue Form der Rechtspsychologie (*law and psychology*) charakterisiert, die sich nicht auf das Studium einzelner Akteure des Rechtssystems (Geschworene, Augenzeugen und Kriminelle) beschränkt, sondern Prozessparteien, Konsumenten, Produzenten, Vertragsparteien, Gesellschaftsorgane, Eltern und Arbeitnehmer in ihr Blickfeld einbezieht.²⁹⁸ Das Recht hat sich historisch auf „Modelle“ menschlichen Verhaltens gestützt, die anekdotisch, introspektiv und ideologisch begründet wurden. Die Ökonomische Analyse des Rechts hat dem eine kohärente, ausformulierte Theorie menschlichen Verhaltens gegenüberstellt, die aber keine empirische Basis hat. Die moderne Psychologie bietet wissenschaftliche, empirisch abgestützte Theorien menschlichen Verhaltens. Die Rechtswissenschaft wird profitieren, wenn sie sich diese Erkenntnisse vermehrt zu Nutzen macht.²⁹⁹ Damit wird auch klar, dass die *Behavioral Law and Economics*, trotz ihres trügerischen Namens, auch für Juristen und Rechtssysteme interessant sind, die mit der Ökonomischen Analyse des Rechts nichts am Hut haben. Die im Hauptteil dieser Arbeit vorgestellten Auswirkungen kognitiver Täuschungen auf das Urteil von Richtern und Parteien in Zivil- und Strafverfahren sind auf jeden Fall relevant, auch wenn man Effizienz nicht als Rechtsziel ansieht.
- 124 Der Mangel an einer allgemeingültigen Theorie zwingt Juristen, die psychologische Erkenntnisse für die Analyse juristischer Phänomene beiziehen, den sprichwörtlichen Lehnstuhl zu verlassen und vermehrt selber empirisch tätig zu werden. Schwache Arbeiten der BLE stellen in einer kurzen Einleitung eine in psychologischen Experimenten beobachtete Verhaltensweise vor und analysieren dann ein Rechtsgebiet unter der Annahme, dass sich alle Rechtsunterworfenen gemäss den experimentellen Befunden verhalten.³⁰⁰ Überzeugende Arbeiten gehen ebenfalls von den experimentellen Befunden aus, überprüfen diese aber anhand eigener Experimente und, wenn vorhanden, Daten aus dem Feld.³⁰¹

²⁹⁸ KOROBKIN/ULEN, FN 104, 766.

²⁹⁹ KOROBKIN/ULEN, FN 104, 766.

³⁰⁰ MITCHELL, FN 200, 71 ff.

³⁰¹ RACHLINSKI, FN 219, ist ein Beispiel für eine überzeugende Arbeit.

§ 6 Prospect Theory – ein deskriptives Modell menschlichen Risikoverhaltens

- 125 Als verbreiteste formalisierte Alternative zur Erwartungsnutzentheorie wird hier die Prospect Theory ausführlicher dargestellt. Da sie strukturell vergleichbar ist mit der Erwartungsnutzentheorie bietet sie für den Leser, der mit der Erwartungsnutzentheorie nicht vertraut ist, auch eine Einführung in die methodischen Grundlagen dieser Theorie. Anders als die Erwartungsnutzentheorie ist die Prospect Theory jedoch eine deskriptive Theorie menschlichen Risikoverhaltens; sie sagt nicht, wie man sich verhalten soll, sondern beschreibt, wie sich Menschen tatsächlich verhalten. Obwohl die Prospect Theory mittels hypothetischer Lotterien entwickelt wurde,³⁰² sind ihre Voraussagen keineswegs nur auf das Verhalten unter Laborbedingungen beschränkt.³⁰³ Sie wurden bestätigt, wenn es für die Versuchspersonen um Einsätze in der Höhe mehrerer Monatslöhne ging;³⁰⁴ und auch Experten sind nicht gegen den Einfluss des Darstellungseffekts gefeit.³⁰⁵ Die Voraussagen der Prospect Theory bezüglich spezifisch juristischer Entscheidungen, wie diejenige, einen Vergleich abzuschliessen, eine Vertragsklausel zu akzeptieren oder Steuern zu hinterziehen, werden hinten, S. 90 ff., dargestellt.

I. Ein Gedankenexperiment zum Einstieg

- 126 Angenommen, sie haben die Wahl zwischen den folgenden beiden Optionen:
- a) ein sicherer Gewinn von Fr. 240;
 - b) eine Chance von 25 %, Fr. 1'000 zu gewinnen und eine Chance von 75%, nichts zu gewinnen.
- 127 Welche wählen sie? Wenn sie sich wie die grosse Mehrheit entscheiden, haben sie den sicheren Gewinn gewählt (84 % aller Versuchspersonen in einem Experiment von AMOS TVERSKY und DANIEL KAHNEMAN wählten die Option a),³⁰⁶ sie zeigen somit eine Ab-

³⁰² Eine Zusammenfassung zahlreicher Laborexperimente bieten IRWIN P. LEVIN/SANDRA L. SCHNEIDER/GARY J. GAETH, All Frames Are not Created Equal: A Typology and Critical Analysis of Framing Effects, *Organizational Behavior and Human Decision Processes* 1998, 149-188.

³⁰³ COLIN F. CAMERER, Prospect Theory in the Wild: Evidence from the Field, in: KAHNEMAN/TVERSKY (Hrsg.), 288-300; LORENZ GOETTE/ERNST FEHR/DAVID HUFFMANN, Loss Aversion and Labor Supply, *Journal of the European Economic Association* 2004, 216-228; SHLOMO BENARTZI/RICHARD H. THALER, Myopic Loss Aversion and the Equity Premium Puzzle, *Quarterly Journal of Economics* 1995, 73-92; ERIC J. JOHNSON/JOHN HERSHEY/JACQUELINE MEZAROS/HOWARD KUNREUTHER, Framing, Probability Distortions, and Insurance Decisions, *Journal of Risk and Uncertainty* 1993, 35-51. Die beiden letztgenannten Aufsätze sind ebenfalls in dem in der Aufsatzsammlung KAHNEMAN/TVERSKY (Hrsg.) abgedruckt.

³⁰⁴ STEVEN J. KACHELMEIER/MOHAMED SHEHATA, Examining Risk Preferences Under High Monetary Incentives: Experimental Evidence From the People's Republic of China, *American Economic Review* 1992, 1120-1141.

³⁰⁵ WILLIAM J. QUALLS/CHRISTOPHER P. PUTO, Organizational Climate and Decision Framing: An Integrated Approach to Analyzing Industrial Buying Decisions, *Journal of Marketing Research* 1989, 179-192 (für Einkaufsmanager); MICHAEL J. ROSZKOWSKI/GLEN E. SNELBECKER, Effects of "Framing" on Measures of Risk Tolerance: Financial Planners are Not Immune, *Journal of Behavioral Economics* 1990, 237-246.

³⁰⁶ AMOS TVERSKY/DANIEL KAHNEMAN, The Framing of Decisions and the Psychology of Choice, *Science* 1981, 453-458, 456.

neigung gegenüber der riskanten Wahl, wenn es um Gewinne geht (obwohl die riskante Wahl einen höheren erwarteten Wert von $0,25 \cdot \text{Fr. 1'000} = \text{Fr. 250}$ hat).

128 Welche Option wählen sie, wenn die folgenden beiden Alternativen zur Auswahl stehen:

- c) ein sicherer Verlust von Fr. 750;
- d) eine Chance von 75 %, Fr. 1'000 zu verlieren und eine Chance von 25%, nichts zu verlieren.

129 Wenn sie wiederum wie die überwiegende Mehrheit entscheiden, haben sie in diesem Fall nicht den sicheren Verlust, sondern die riskante Option d) gewählt (87 % aller Versuchspersonen entschieden sich für diese Wahl);³⁰⁷ sie verhalten sich somit risikogeneigt, wenn es um die Wahl zwischen Verlusten geht. Dieses Verhalten widerspricht der klassischen ökonomischen Theorie, gemäss der das Risikoverhalten einzig vom Verhältnis des Vermögens einer Person zur Grösse der Gewinn- oder Verlustchance abhängen sollte.³⁰⁸

130 Das beobachtete Verhalten ist nicht auf Spiele mit finanziellen Einsätzen beschränkt und lässt sich selbst dann beobachten, wenn die Ergebnisse des „Spiels“ nur anders dargestellt (*framed*) werden, aber objektiv identisch sind, wie der folgende Klassiker aus der *Framing* Literatur zeigt:³⁰⁹

Stellen sie sich vor, dass sich die USA auf den Ausbruch einer ungewöhnlichen asiatischen Krankheit vorbereiten, von der erwartet wird, dass 600 Personen daran sterben werden. Es wurden zwei verschiedene Pläne vorgeschlagen, die Krankheit zu bekämpfen. Nehmen sie an, dass die Folgen der beiden Pläne genau bekannt sind:

- Wenn Plan A umgesetzt wird, werden 200 Personen gerettet.
- Wenn Plan B umgesetzt wird, besteht eine Wahrscheinlichkeit von einem Drittel ($1/3$), dass 600 Personen gerettet werden, und eine Wahrscheinlichkeit von zwei Dritteln ($2/3$), dass niemand gerettet wird.

131 72 % der Versuchspersonen wählten den Plan A, der 200 Menschen sicher rettet;³¹⁰ wenn beide Alternativen in der Form von „Gewinnen“ dargestellt werden, entscheidet sich daher eine Mehrheit für die risikolose Wahl. Einer weiteren Gruppe von Versuchspersonen wurde der gleiche Sachverhalt geschildert, aber ihnen wurden die folgenden zwei Pläne zur Auswahl angeboten:

- Wenn Plan C umgesetzt wird, werden 400 Personen sterben.

³⁰⁷ TVERSKY/KAHNEMAN, FN 306, 454.

³⁰⁸ Siehe beispielsweise HARRY MARKOWITZ, The Utility of Wealth, Journal of Political Economy 1952, 151-158 (erhältlich unter <http://cowles.econ.yale.edu/P/cp/p00b/p0057.pdf>, besucht am 23. März 2005).

³⁰⁹ Eine Meta-Analyse des „asiatische Krankheit Problems“ bieten ANTON KÜHLBERGER/MICHAEL SCHULTE-MECKLENBECK/JOSEF PERNER, The Effects of Framing, Reflection, Probability, and Payoff on Risk Preference in Choice Tasks, Organizational Behavior and Human Decision Processes 1999, 204-231. Das asiatische Krankheit Problem reagiert sensibel auf Änderungen des Wortlauts der Fragen; KÜHLBERGER et al. kommen aber zum Schluss, dass die Darstellung die von der Prospect Theory vorausgesagten Auswirkungen auf das Risikoverhalten hat.

³¹⁰ TVERSKY/KAHNEMAN, FN 306, 453.

- Wenn Plan D umgesetzt wird, besteht eine Wahrscheinlichkeit von einem Drittel (1/3), dass niemand sterben wird, und eine Wahrscheinlichkeit von zwei Dritteln (2/3), dass 600 Menschen sterben werden.

132 In diesem Fall haben 78 % der Versuchspersonen den (riskanten) Plan D gewählt.³¹¹ Die Pläne A und C und die Pläne B und D sind offensichtlich identisch. Ihre respektiven Folgen wurden nur anders dargestellt – bei den Plänen A und B als „Gewinne“ und bei den Plänen C und D als „Verluste“. Wie KAHNEMAN und TVERSKY festgestellt haben, hat die Darstellung (*framing*) der Alternativen als Gewinne oder Verluste einen massgeblichen Einfluss auf das Risikoverhalten; man spricht daher vom Darstellungseffekt.

A. Das Allais Paradox

133 Der französische Ökonom MAURICE ALLAIS wies bereits 1953 darauf hin, dass Menschen sich nicht gemäss dem (kurz vorher von VON NEUMANN und MORGENSTERN zu einem axiomatischen Modell formalisierten) Erwartungsnutzenmodell verhalten.³¹² KAHNEMAN und TVERSKY haben ihren Versuchspersonen folgende Variationen des ALLAIS-Paradox (gleicher Aufbau, aber weniger extreme Zahlen als ALLAIS verwendete) zur Auswahl angeboten:³¹³

- a) Fr. 2'400 auf sicher
- b) 33 % Chance, Fr. 2'500 zu gewinnen; und
66 % Chance, Fr. 2'400 zu gewinnen; und
1 % Chance, nichts zu gewinnen.

134 82 % der Versuchspersonen wählen Option a) bei der obigen Wahl. Hingegen ziehen 83 % Option d) vor, wenn sie die Wahl zwischen den folgenden Alternativen haben:³¹⁴

- c) 34 % Chance, Fr. 2'400 zu gewinnen; und
66 % Chance, nichts zu gewinnen.
- d) 33 % Chance, Fr. 2'500 zu gewinnen; und
67 % Chance, nichts zu gewinnen.

135 Dies verletzt aber das Unabhängigkeits-Axiom der Erwartungsnutzen-Theorie. Wer Alternative a) der Alternative b) vorzieht, für den gilt

$$1 \cdot u(\text{Fr. 2'400}) > 0,33 \cdot u(\text{Fr. 2'500}) + 0,66 \cdot u(\text{Fr. 2'400}) + 0,01 \cdot u(\text{Fr. 0}) \quad (2)$$

was sich vereinfachen lässt zu

³¹¹ TVERSKY/KAHNEMAN, FN 306, 453.

³¹² MAURICE ALLAIS, Le comportement de l'homme rationel devant le risque, critique des postulats et axiomes de l'école Americaine, *Econometrica* 1953, 503-546.

³¹³ DANIEL KAHNEMAN/AMOS TVERSKY, Prospect theory: An analysis of decision under risk, *Econometrica* 1979, 263-291; Nachdruck in: KAHNEMAN/TVERSKY (Hrsg.), 17-43, 20.

³¹⁴ KAHNEMAN/TVERSKY, FN 313, 20.

$$0,34 \cdot u(\text{Fr. 2'400}) > 0,33 \cdot u(\text{Fr. 2'500}) \quad (3)$$

136 Wer bei der zweiten Wahl d) vorzieht, für den gilt

$$0,33 \cdot u(\text{Fr. 2'500}) > 0,34 \cdot u(\text{Fr. 2'400}) \quad (4)$$

und beide Ungleichungen (3) und (4) können offensichtlich nicht gleichzeitig zutreffen, ganz egal, wie die persönliche Nutzenfunktion u aussieht.³¹⁵

137 Die Verletzung des Unabhängigkeits-Axioms wird ersichtlich, wenn man das ALLAIS-Paradox in der erstmals von LEONHARD SAVAGE vorgeschlagenen Art und Weise darstellt.³¹⁶ Die verschiedenen Ausgänge der Alternativen werden als verschiedenfarbige Bälle konzeptualisiert, die aus einer Urne mit 100 Bällen gezogen werden. Bei der Alternative A erhält der Wählende Fr. 2'400, wenn einer der 66 schwarzen, der 33 weissen oder der eine rote Ball gezogen wird. Bei der Alternative B erhält er Fr. 2'400, wenn einer der 66 schwarzen Bälle gezogen wird und Fr. 2'500, wenn einer der 33 weissen Bälle gezogen wird, aber nichts, wenn der rote Ball gezogen wird. Wie aus der Tabelle 5 leicht ersichtlich ist, ist der Ausgang des Spiels bei der ersten Wahl, wenn ein schwarzer Ball gezogen wird, derselbe unabhängig davon, ob der Wählende die Alternative A oder B gewählt hat. Dasselbe gilt bei der zweiten Wahl. Es kommt daher nur auf die mit der Ziehung eines weissen oder des roten Balles verbundenen Ausgänge an. Da diese Ausgänge bei A und C und bei B und D identisch sind, sollte, wer bei der ersten Wahl A wählt, bei der zweiten Wahl C wählen. Wer A und D wählt, bevorzugt in einem Fall $0,34 \cdot \text{Fr. 2'400}$, im anderen Fall $0,33 \cdot \text{Fr. 2'500}$.

		Bälle in der Urne		
		66 schwarze Bälle	33 weisse Bälle	1 roter Ball
Erste Wahl	A	Fr. 2'400	Fr. 2'400	Fr. 2'400
	B	Fr. 2'400	Fr. 2'500	nichts
Zweite Wahl	C	Nichts	Fr. 2'400	Fr. 2'400
	D	Nichts	Fr. 2'500	nichts

Tabelle 5: Darstellung des Allais-Paradoxes nach SAVAGE (1954)

II. Prospect Theory

A. Referenzpunkt

138 Wie vorne, S. 30, erwähnt, geht das Erwartungsnutzenmodell vom Vermögensstand als Träger des Nutzens aus. Die Prospect Theory hingegen nimmt an, dass die *Veränderung*, ausgehend von einem Referenzpunkt, entscheidend ist. Gemäss Prospect Theory wird der Ausgang eines riskanten Spiels, ausgehend von einem Referenzpunkt, als Gewinn oder

³¹⁵ Natürlich lassen sich persönliche Nutzenfunktionen konstruieren, die das Resultat erklären, aber solche Konstruktionen erscheinen immer post hoc.

³¹⁶ LEONHARD J. SAVAGE, *The Foundations of Statistics*, New York 1954, 103 ff.

Verlust kodiert, wobei die Nutzenfunktion für Gewinne und Verluste unterschiedlich ist.³¹⁷ Welcher Zustand als Referenzpunkt genommen wird, lässt die Prospect Theory dabei offen (was ihr auch Kritik eingetragen hat). In den meisten Fällen ist jedoch mit hinreichender Sicherheit klar, welches der Referenzpunkt ist. Am häufigsten handelt es sich dabei um den Ist-Zustand, den Status quo. Der andere wichtige Referenzpunkt, den Menschen für die Kodierung eines Ergebnisses als Gewinn oder Verlust verwenden, ist der Soll-Zustand (*aspiration level*).³¹⁸ Ein Anwalt kann sich beispielsweise vornehmen, einen Jahresumsatz von mindestens Fr. x zu erzielen; alles, was er weniger an Umsatz erzielt hat, wird als Verlust kodiert, ein Mehrumsatz über das Erhoffte hinaus als Gewinn. Die meisten Leute neigen dazu, Risiken zu vermeiden, wenn sie zwischen Optionen wählen, die Gewinne darstellen, tendieren aber dazu, Risiken einzugehen, um Verluste zu vermeiden.³¹⁹

B. Wertfunktion

- 139 Die Prospect Theory schlägt zur Abbildung dieses Risikoverhaltens eine S-förmige Wertfunktion (*value function*) vor, die am Referenzpunkt einen Knick aufweist und für Gewinne konkav und für Verluste konvex verläuft (siehe Abbildung 6). Die Krümmung bildet das psychophysikalische Prinzip ab, dass der Unterschied zwischen 0 und 100 subjektiv als grösser empfunden wird als der Unterschied zwischen 1'000 und 1'100.³²⁰ Der Knick am Referenzpunkt bedeutet, dass Menschen sich im Bereich möglicher Gewinne risikoscheu (*risk averse*), im Bereich möglicher Verluste aber risikogeneigt (*risk seeking*) verhalten.

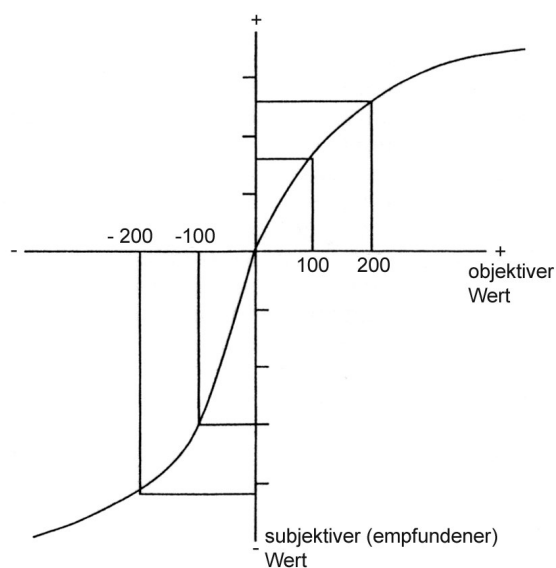


Abbildung 6: Hypothetische Wertfunktion der Prospect Theory

³¹⁷ DANIEL KAHNEMAN/AMOS TVERSKY, Choices, Values, and Frames, American Psychologist 1984, 341-350; Nachdruck in: KAHNEMAN/TVERSKY (Hrsg.), 1-16, 3.

³¹⁸ TVERSKY/KAHNEMAN, FN 306, 456; CHIP HEATH/RICHARD P. LARRICK/GEORGE WU, Goals as Reference Points, Cognitive Psychology 1999, 79-109.

³¹⁹ KAHNEMAN/TVERSKY, FN 317, 3.

³²⁰ TVERSKY/KAHNEMAN, FN 306, 454.

- 140 KAHNEMAN und TVERSKY haben ihren Versuchspersonen hunderte von hypothetischen Spielen der Art des Beispiels am Anfang dieses Kapitels angeboten und die Eigenschaften der Wertfunktion empirisch bestimmt.³²¹ Eine typische Wertfunktion – es gibt individuelle Unterschiede, nicht jeder Mensch hat genau dieselbe Wertfunktion – lässt sich durch folgende Funktion v abbilden:

$$v(x) = \begin{cases} x^\alpha & \text{wenn } x > 0 \\ -\lambda (-x^\alpha) & \text{wenn } x < 0 \end{cases} \quad (\text{mit typischen Werten von } \alpha \approx 0,88 \text{ und } \lambda \approx 2,25)^{322} \quad (5)$$

1. Loss aversion

- 141 Eine weitere Eigenschaft der Wertfunktion der Prospect Theory ist, dass sie im Bereich der Verluste steiler verläuft als im Bereich der Gewinne. Der „Unwert“ eines Verlustes $-X_1$ wird rund doppelt so stark empfunden wie der Wert des gleichen Gewinns X_1 (daher $\lambda \approx 2,25$).³²³ Die verstärkte Abneigung gegenüber Verlusten wird als „*loss aversion*“ bezeichnet.³²⁴ *Loss aversion* kann den *Status quo Bias* erklären. Als *Status quo Bias* wird die Neigung bezeichnet, eine Wahl zu treffen, die den Ist-Zustand erhält oder leicht verbessert.³²⁵
- 142 *Loss aversion* wird auch bemüht, um den so genannten Besitztumseffekt (*endowment effect*) zu erklären. In einem typischen Experiment zur Demonstration des Besitztumseffekts erhält die Hälfte der Versuchspersonen ein Gut – beispielsweise ein Becher mit dem Logo der Universität – während die andere Hälfte der Versuchspersonen nichts erhält. Diejenigen Versuchspersonen, die nichts erhalten haben, werden gefragt, wie viel sie für einen Becher der Art, wie sie die andere Gruppe der Versuchspersonen erhalten hat, zu zahlen bereit wären (Bereitschaft zu kaufen). Die Versuchspersonen, die einen Becher besitzen (und die sie behalten dürfen), werden gefragt, zu welchem Preis sie den Becher verkaufen würden (Bereitschaft zu verkaufen). Gemäss traditioneller ökonomischer Analyse müssen die Preise der Käufer und Verkäufer im Schnitt gleich sein, denn die Präferenzen für das Gut hängen nicht vom Besitz des Gutes ab und sind zufällig verteilt. Da die Versuchspersonen ebenfalls zufällig auf die beiden Gruppen verteilt werden, sollte es keinen signifikanten Unterschied geben. Zahllose Experimente aus über 30 Jahren mit Gütern wie Schokoriegeln, Schreibstiften, Bechern, Kinotickets, Jagdpatenten, Atommülldeponien, übel riechenden Flüssigkeiten, möglicherweise verseuchten Sandwich und anderen zeigen

³²¹ AMOS TVERSKY/DANIEL KAHNEMAN, *Advances in Prospect Theory: Cumulative Representation of Uncertainty*, Journal of Risk and Uncertainty 1992, 297-323; Nachdruck in: KAHNEMAN/TVERSKY (Hrsg.), 44-65, 55.

³²² TVERSKY/KAHNEMAN, FN 321, 59.

³²³ AMOS TVERSKY/DANIEL KAHNEMAN, *Loss Aversion in Riskless Choice: A Reference-Dependent Model*, Quarterly Journal of Economics 1991, 1039-1061; DANIEL KAHNEMAN/JACK L. KNETSCH/RICHARD H. THALER, *Experimental Tests of the Endowment Effect and the Coarse Theorem*, Journal of Political Economy 1990, 1325-1348; Nachdruck in: SUNSTEIN (Hrsg.), 211-231.

³²⁴ DANIEL KAHNEMAN/JACK L. KNETSCH/RICHARD H. THALER, *The Endowment Effect, Loss Aversion and the Status Quo Bias*, Journal of Economic Perspectives 1991, 193-206.

³²⁵ AMOS TVERSKY/DANIEL KAHNEMAN, *Rational Choice and the Framing of Decisions*, Journal of Business 1986, 251-278, 257.

aber,³²⁶ dass der verlangte Verkaufspreis rund doppelt so hoch ist wie der angebotene Kaufpreis; dies gilt auch dann, wenn keine Anreize für strategisches über- resp. untertreiben des deklarierten Preises bestehen und die Versuchspersonen dies verstanden haben.³²⁷ Da der beobachtete Unterschied zwischen offeriertem Kaufpreis und verlangtem Verkaufspreis einzig vom Besitz des Gutes abzuhängen scheint, wird er als „Besitztumseffekt“ bezeichnet.³²⁸

- 143 *Loss aversion* kann den „sofortigen Besitztumseffekt“³²⁹ „erklären“ – in Anführungszeichen deshalb, weil damit der Besitztumseffekt letztlich nur auf ein allgemeineres Phänomen zurückgeführt wird, das ohne psychologische Begründung bleibt.³³⁰ die Versuchspersonen, die das Gut erhalten haben, betrachten die Aufgabe des Besitzes als Verlust verglichen mit dem Status quo. Dagegen betrachten die Versuchspersonen, die kein Gut erhalten haben, den Erwerb des Gutes, verglichen mit dem Status quo, als Gewinn. Da ein Verlust gemäss der Wertfunktion der Prospect Theory rund doppelt so stark empfunden wird wie ein gleich grosser Gewinn, verlangen die besitzenden Versuchspersonen rund zwei Mal mehr für die Aufgabe des Gutes, als die nicht-besitzenden zu zahlen bereit sind (für eine weitere Erklärung des Besitztumseffekts siehe hinten, S. 102 ff.).
- 144 *Loss aversion* vermag auch zu erklären, warum die meisten Deutschen nie bereit wären, für einen höheren Lohn auf ihren gesetzlichen Urlaubsanspruch (von mindestens 24 Werktagen, § 3 Bundesurlaubsgesetz) zu verzichten, während Amerikaner, die in der Regel maximal zwei Wochen Urlaub beziehen, nicht bereit sind, weniger zu verdienen, um mehr Urlaub zu bekommen. Für die Deutschen sind 24 Tage Urlaub der Status quo, die Aufgabe von Urlaubstagen wird daher als Verlust kodiert, während für die Amerikaner die zwei Wochen Urlaub der Status quo sind und mehr Urlaub ein Gewinn darstellt.
- 145 Warum Verluste doppelt so stark empfunden werden wie korrespondierende Gewinne, ist bis heute nicht geklärt. Vertreter der evolutionären Psychologie vermuten, dass Verluste manchmal den Tod und damit die Nichtweitergabe der eigenen Gene bedeuteten, während Gewinne die Wahrscheinlichkeit, die eigenen Gene der nächsten Generation zu vererben, nicht im gleichen Mass erhöhten.³³¹ Eine andere evolutionspsychologische Erklärung ist, dass die Abneigung gegen Verluste eine Form territorialer Verteidigung ist. Biologische Studien zeigen, dass Tiere viel härter kämpfen, um ein Revier zu verteidigen, als um es zu erobern. Dies ergibt wahrscheinlich Sinn, denn das besitzende Tier hat einen grösseren Nutzen davon, in einem Revier zu bleiben, das es bereits gut kennt (der Eindringling muss die Kosten der Informationssuche erst investieren, weshalb der Wert des Reviers für ihn

³²⁶ JOHN K. HOROWITZ/KENNETH E. MCCONNELL, A Review of WTA/WTP Studies, *Journal of Environmental Economics and Management* 2002, 426-447.

³²⁷ KAHNEMAN/KNETSCH/THALER, FN 324, 231 ff.

³²⁸ THALER, FN 163 39-60.

³²⁹ TVERSKY/KAHNEMAN, FN 323, 1041.

³³⁰ Einen Überblick über die zahlreichen psychologischen und ökonomischen Erklärungsansätze für den beobachteten Unterschied zwischen Kauf- und Verkaufspreis bietet SCHWEIZER, FN 19.

³³¹ ELIZABETH HOFFMAN/MATTHEW L. SPITZER, Willingness to Pay vs. Willingness to Accept: Legal and Economic Implications, *Washington University Law Quarterly* 1993, 59-114, 89.

entsprechend geringer ist).³³² Diese Erklärungen sind faszinierend, sind aber m.E. wie viele der evolutionspsychologischen „just so“ Geschichten mit Vorsicht zu geniessen.

C. Gewichtungsfunktion

- 146 Neben der Risikoneigung im Bereich der Verluste haben KAHNEMAN und TVERSKY noch ein anderes Verhalten ihrer Versuchspersonen festgestellt: wenn ein grosser Gewinn sehr unwahrscheinlich ist, verhalten sich Menschen auch im Bereich der Gewinne risikogeneigt. Die meisten Menschen ziehen die Option B vor, wenn die Wahl besteht zwischen A = 45 % Chance, Fr. 6'000 zu gewinnen und B = 90 % Chance, Fr. 3'000 zu gewinnen.³³³ Dies ist ein typischer Fall der Risikoscheu im Bereich der Gewinne; man zieht den (fast) sicheren kleineren Gewinn einem potenziellen grösseren Gewinn vor („Lieber den Spatz in der Hand als die Taube auf dem Dach“). Hingegen ziehen 73 % der Versuchspersonen die Option C vor, wenn die Wahl besteht zwischen C = 1 Promille Chance, Fr. 6'000 zu gewinnen oder D = 2 Promille Chance, Fr. 3'000 zu gewinnen.³³⁴ Auch im zweiten Fall ist die Chance, Fr. 6'000 zu gewinnen, genau halb so gross wie die Chance, Fr. 3'000 zu gewinnen. Trotzdem scheint in diesem Fall – wenn beide Chancen sehr gering sind – die riskante Wahl attraktiver. Die Vorstellung, Fr. 6'000 gewinnen zu können, lässt die geringe Chance verbllassen.
- 147 Um dieses Verhalten ebenfalls durch die Prospect Theory beschreiben zu können, schlagen KAHNEMAN und TVERSKY eine Gewichtungsfunktion (*weighing function*) der objektiven Wahrscheinlichkeit vor (die ursprüngliche Prospect Theory befasste sich nur mit riskanten Wahlen, d.h. mit der Entscheidung zwischen Spielen, bei denen die relativen Häufigkeiten der Ergebnisse bekannt sind). Die Gewichtungsfunktion hat typischerweise – auch hier gibt es wieder individuelle Unterschiede – die Form gemäss Abbildung 7.

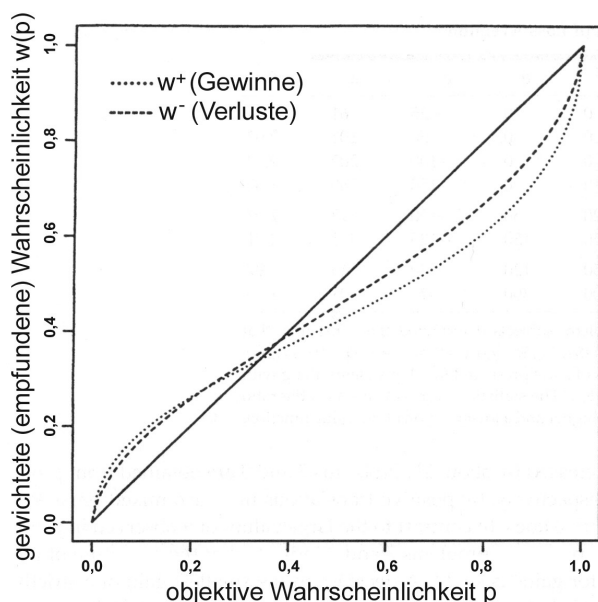


Abbildung 7: Gewichtungsfunktion der Prospect Theory (aus TVERSKY/KAHNEMAN, 1992/2000)

³³² OWEN D. JONES, Time-Shifted Rationality and the Law of Law's Leverage: Behavioral Economics Meets Behavioral Biology, Northwestern University Law Review 2001, 1141-1205, 1185.

³³³ 86 % ziehen B vor, KAHNEMAN/TVERSKY, FN 313, 21 f.

³³⁴ KAHNEMAN/TVERSKY, FN 313, 21 f.

148 Die Gewichtungsfunktion hat folgende Eigenschaften:

- a) sie verläuft im Bereich der Endpunkte sehr steil, was bedeutet, dass Menschen sehr sensibel auf eine Veränderung der Wahrscheinlichkeit in diesem Bereich reagieren. Menschen sind beispielsweise bereit, mehr Geld auszugeben, ein Risiko von 0,001 auf 0 zu reduzieren als von 0,2 auf 0,15, obwohl die Abnahme des Risikos im zweiten Fall objektiv grösser ist. Das subjektive Empfinden kann dazu führen, dass Geld in Sicherheitsmassnahmen investiert wird, das an einem anderen Ort objektiv viel mehr zur Sicherheit der Menschen beitragen könnte.³³⁵
- b) geringe objektive Wahrscheinlichkeiten (unter ca. 0,3) werden tendenziell überschätzt. Bei ungefähr 0,3-0,4 sind die Leute gut kalibriert, d.h. objektive Wahrscheinlichkeit und subjektives Empfinden stimmen überein.
- c) mittlere und hohe Wahrscheinlichkeiten – über ca. 0,4 – werden tendenziell unterschätzt. Gleichzeitig verläuft die Kurve im Bereich zwischen ca. 0,5 und 0,9 zu flach; d.h. wesentliche Änderungen der objektiven Wahrscheinlichkeit in diesem Bereich bewirken eine nur geringe Änderung des Entscheidungsgewichts.
- d) im Bereich sehr hoher Wahrscheinlichkeiten verläuft die Kurve wieder sehr steil, was abbildet, dass es subjektiv einen grossen Unterschied macht, Fr. 1'000 mit einer Wahrscheinlichkeit von 0,99 zu gewinnen oder mit Sicherheit zu bekommen („certainty effect“).³³⁶
- e) die Gewichtungsfunktion verläuft zudem für Gewinne und Verluste leicht unterschiedlich (siehe Abbildung 7); aber für die (groben) Voraussagen, die aufgrund der Prospect Theory im Bereich des Rechts getroffen werden können, spielt dieser Unterschied keine wesentliche Rolle.

149 Die Gewichtungsfunktion π lässt sich durch folgende Formel beschreiben³³⁷

$$\pi(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{1/\gamma}} \quad \text{wobei } \gamma \approx 0,61 \text{ für Gewinne } (\approx 0,69 \text{ für Verluste}) \quad (6)$$

D. Kombination von Wert- und Gewichtungsfunktion

150 Die Prospect Theory postuliert, dass der subjektive Wert V eines Prospects, d.h. einer Wahl zwischen riskanten Alternativen, aus objektivem Wert und objektiver Wahrscheinlichkeit resultiert, die jeweils von der Wert- resp. Gewichtungsfunktion der Prospect Theory „verzerrt“ werden. Es gilt also

$$V(x,p) = v(x) \cdot \pi(p) \quad (7)$$

³³⁵ Beispiel bei HASTIE/DAWES, FN 399, 297.

³³⁶ KAHNEMAN/TVERSKY, FN 317, 7 ff.

³³⁷ TVERSKY/KAHNEMAN, FN 321, 58.

- 151 Aus der Kombination von Wert- und Gewichtungsfunktion folgt das typische vierteilige Risikoverhalten (*four fold risk pattern*) der Prospect Theory: Risikoscheu im Bereich der Gewinne mit mittleren und hohen Wahrscheinlichkeiten und der Verluste mit geringen Wahrscheinlichkeiten; risikogeneigt im Bereich der Verluste mit mittleren und hohen Wahrscheinlichkeiten und der Gewinne mit kleinen Wahrscheinlichkeiten (siehe Tabelle 6).

	Geringe Wahrscheinlichkeit	Mittlere bis hohe Wahrscheinlichkeit
Gewinne	risikogeneigt	risikoscheu
Verluste	risikoscheu	risikogeneigt

Tabelle 6: Risikoverhalten gemäss Prospect Theory

E. Die Unterscheidung zwischen Risiko und Unsicherheit

- 152 In der Entscheidungstheorie wird zwischen Entscheidungen unter Risiko und Entscheidungen unter Unsicherheit unterschieden.³³⁸ Entscheidungen unter Risiko sind Entscheidungen, bei denen die relative Häufigkeit der Ergebnisse des „Spiels“ bekannt sind – typische Beispiele sind Glücksspiele, wie der Wurf einer fairen Münze oder das Rollen eines fairen Würfels. Eine Wahl ist hingegen unsicher, wenn die objektive Wahrscheinlichkeit ihrer Ergebnisse nicht bekannt sind – zum Beispiel ist der Ausgang der Entscheidung, eine Klage einzureichen, unsicher, denn die Wahrscheinlichkeit, zu obsiegen, lässt sich nicht mathematisch berechnen. Juristische Entscheidungen sind typischerweise Entscheidungen unter Unsicherheit, nicht unter Risiko. Die ursprüngliche Prospect Theory (1979) beschränkt sich explizit auf die Modellierung riskanter, nicht unsicherer, Wahlen.³³⁹ Menschen messen einem Ereignis aber oft eine subjektive Wahrscheinlichkeit bei, selbst wenn sich die Wahrscheinlichkeit des Eintretens – beispielsweise bei einem singulären Ereignis – nicht in objektiven Wahrscheinlichkeiten ausdrücken lässt (näheres zur Unterscheidung von subjektiver und objektiver Wahrscheinlichkeit hinten, S. 140 ff.). Klienten verlangen vom Anwalt geradezu eine solche Einschätzung der Wahrscheinlichkeit des Obsiegens (und Anwälte winden sich, sich gegenüber dem Klienten auf eine Prozentzahl festzulegen. Trotzdem hat jeder Anwalt eine – allerdings systematisch verzerrte, siehe S. 265 – Ahnung, wie der Fall ausgehen wird). Es hat sich gezeigt, dass die Prospect Theory auch den Umgang mit subjektiven Wahrscheinlichkeiten gut abbilden kann; die erweiterte Fassung der Prospect Theory (1992) erlaubt denn auch die Anwendung der Prospect Theory auf Entscheidungen unter Unsicherheit.³⁴⁰ Die Prospect Theory vermag daher auch juristische Entscheidungen, wie z.B. den Entscheid, Klage einzureichen, zu erklären.

³³⁸ FRANK H. KNIGHT, Risk, Uncertainty and Profit, Boston 1921, 21.

³³⁹ KAHNEMAN/TVERSKY, FN 313, 18.

³⁴⁰ TVERSKY/KAHNEMAN, FN 321, 45.

F. Anwendungsbeispiele

- 153 Die Alternativen, die beim Problem der asiatischen Krankheit (S. 54) zur Wahl stehen, lassen sich mittels Prospect Theory wie folgt modellieren:

$$\begin{aligned}
 \text{a) } V(+200; 1) &= \pi(1) \cdot v(+200) &= 105,9 \\
 \text{b) } V(+600; 0,33; 0; 0,67) &= \pi(0,33) \cdot v(+600) + \pi(0,67) \cdot v(0) &= 96,9 \\
 \text{c) } V(+400; 1) &= \pi(1) \cdot v(-400) &= -438,5 \\
 \text{d) } V(-600; 0,67; 0; 0,33) &= \pi(0,67) \cdot v(-600) + \pi(0,33) \cdot v(0) &= -349,5
 \end{aligned}$$

- 154 Der subjektive Wert des Planes A ist gemäss Prospect Theory höher als der subjektive Wert des Planes B; ebenso ist der subjektive Wert des Planes D höher (weniger negativ) als der Wert des Planes C. Folglich sagt die Prospect Theory voraus, dass die meisten Menschen unter den ersten beiden Alternativen A, unter den beiden zweiten Alternativen aber D wählen. Die „Arbeit“ verrichtet hier in erster Linie die Wertfunktion der Prospect Theory, deren Knick am Referenzpunkt voraussagt, dass sich Menschen im Bereich der Verluste risikogeneigt verhalten.

- 155 Das ALLAIS-Paradox resultiert hingegen in erster Linie aus der Gewichtungsfunktion der Prospect Theory. Die vier Alternativen des ALLAIS-Paradox lassen sich gemäss Prospect Theory wie folgt modellieren:

$$\begin{aligned}
 \text{e) } V(+2'400; 1) &= \pi(1) \cdot v(+2'400) &= 943,2 \\
 \text{f) } V(0; 0,01; +2'500; 0,33; +2'400; 0,66) &= \pi(0,33) \cdot v(+2'500) + \pi(0,66) \cdot v(+2'400) &= 806 \\
 \text{g) } V(+2'400; 0,34; 0; 0,66) &= \pi(0,34) \cdot v(+2'400) &= 319,7 \\
 \text{h) } V(+2'500; 0,33; 0; 0,67) &= \pi(0,33) \cdot v(+2'500) &= 326,6
 \end{aligned}$$

- 156 Die Prospect Theory sagt also voraus, dass die meisten Menschen bei der Wahl zwischen A und B die Option A vorziehen; bei der Wahl zwischen C und D aber die Option D.

EINZELNE KOGNITIVE TÄUSCHUNGEN UND IHR EINFLUSS AUF DAS URTEIL VON RICHTERN UND PARTEIEN IN GE- RICHTSVERFAHREN

§ 7 Methode, Auswertung, interne und externe Validität der Studie

I. Methode

- 157 Es wurden zwei briefliche Umfragen durchgeführt; die erste im Oktober 2003, die zweite im September 2004. Die Fragebogen sind in den Anhängen B (2003) und C (2004) wiedergegeben. Beide Fragebogen umfassten fünf Sachverhalte, jeder Sachverhalt begann auf einer neuen Seite. Auf der Titelseite wurden die Richterinnen und Richter gebeten, die Fragen ohne Absprache mit ihren Kolleginnen und Kollegen zu beantworten und darauf hingewiesen, dass die Beantwortung etwa eine Viertelstunde in Anspruch nehmen würde (was leicht untertrieben war). Der persönlich adressierte Begleitbrief mit dem Briefkopf des Europäischen Instituts für Rechtspsychologie erläuterte, dass es sich um ein psychologisches Forschungsprojekt handle, ohne genauere Angaben zu machen. Die Umfrage erfolgte anonym; demografische Angaben wurden nicht erhoben. Jedem Fragebogen lag ein adressiertes und frankiertes Rückantwortcouvert bei, und die Richter wurden aufgefordert, den Fragebogen binnen Wochenfrist zu retournieren. Das empfanden einzelne als Zumutung; Erfahrungen aus dem Direkt-Marketing zeigen aber, dass eine kurze Antwortfrist zu einer höheren Rücklaufquote führt.
- 158 Beide Fragebogen wurden einer Voruntersuchung („Pretest“) unterzogen. Als Versuchspersonen des Pretests dienten meine damaligen Arbeitskollegen bei Meyer Lustenberger Rechtsanwälte, Zürich (N ≈ 30). Die Versuchspersonen der Pretests waren alle Juristinnen und Juristen mit zwischen einem und dreissig Jahren Berufserfahrung. Sie sind mit der Testpopulation nicht direkt vergleichbar, da in der Schweiz das Laienrichtertum nach wie vor weit verbreitet ist und die Versuchspersonen des Pretests sich schwerewichtig mit Wirtschaftsrecht befassen, das an erstinstanzlichen Gerichten (wo der grösste Teil der befragten Richter beschäftigt ist) eine geringe Rolle spielt. Trotzdem zeigte der Pretest, welche Sachverhalte angepasst werden mussten und führte zu einigen Änderungen bei der Formulierung der Sachverhalte (wo diese wesentlich sind, werden sie bei den einzelnen Fragen näher erläutert).
- 159 Für die Umfrage vom Herbst 2003 wurden 401 Fragebögen an die Richterinnen und Richter der Zivil- und Strafgerichte der Kantone Aarau, St. Gallen und Zürich verschickt (inklusive der jeweiligen Ober- resp. Kantonsgerichte). Die Fragebogen folgten dem Aufbau der von GUTHRIE, RACHLINSKI und WISTRICH verwendeten Fragebogen, deren Studie die Grundlage für die erste Umfrage bildete.³⁴¹ Da die Fragebogen nicht nur vom Amerikanischen ins Deutsche übersetzt, sondern auch dem schweizerischen Rechtssystem angepasst werden mussten, war eine exakte Replikation ausgeschlossen. Trotzdem ist es interessant, die Resultate der amerikanischen Studie mit denjenigen der vorliegenden

³⁴¹ GUTHRIE/RACHLINSKI/WISTRICH, FN 20, 777. JEFFREY RACHLINSKI stellte mir freundlicherweise die Original-Fragebogen seiner Umfrage zur Verfügung.

Studie zu vergleichen. Jeder Sachverhalt testete eine kognitive Täuschung – Ankereffekt, Darstellungseffekt, Vernachlässigung der Anfangswahrscheinlichkeit, Rückschaufehler und Selbstüberschätzung. Von den Sachverhaltsschilderungen für Ankereffekt, Darstellungseffekt und Rückschaufehler gab es jeweils verschiedene Versionen, um den Einfluss der manipulierten Variable testen zu können. Die Richter wurden auf diesen Umstand nicht hingewiesen. Ich habe keinen Anlass anzunehmen, dass die Richter die Manipulation bemerkt haben. Keine der zahlreichen schriftlichen Reaktionen wies auf diesen Umstand hin, und auch meine Kontakte zu Bezirksrichtern zürcherischer Bezirksgerichte ergaben, dass niemand eine entsprechende Vermutung hegte.³⁴² Die Sachverhaltsvarianten wurden gemischt (wobei die Reihenfolge der Fragen immer gleich blieb). Da es drei Sachverhaltsvarianten für den Ankereffekt, drei Sachverhaltsvarianten für den Rückschaufehler und zwei Varianten für den Darstellungseffekt gab, wurden insgesamt 18 verschiedene Varianten des Fragebogens versandt.

- 160 181, oder 45 %, der angeschriebenen Richter schickten den Fragebogen zumindest teilweise ausgefüllt zurück. Diese – für eine briefliche Umfrage erfreulich hohe – Rücklaufquote entspricht derjenigen, die andere Forscher mit ähnlichen brieflichen Umfragen unter schweizerischen Richtern erzielen konnten.³⁴³ Selbstverständlich bietet das „Non-response-Problem“ Anlass zu methodischer Kritik.³⁴⁴ Die interne Validität (zum Begriff S. 69 f.) wird aber nur dann in Frage gestellt, wenn zwischen Antwortenden und Verweigernden ein systematischer Unterschied besteht. Im vorliegenden Fall muss ein Kritiker postulieren, dass die antwortenden Richter den kognitiven Täuschungen mehr unterliegen als die Verweigerer. Wenn ich auch nicht beweisen kann, dass diese Annahme nicht zutrifft, glaube ich doch, dass sie nicht besonders plausibel ist. Aufgrund einiger Rückmeldungen glaube ich übrigens, dass die antwortenden Richter die Umfrage ernst genommen haben. Einzelne Fragebogen wurden unausgefüllt zurückgeschickt mit der Bemerkung, man entscheide nur über familienrechtliche Streitigkeiten und sehe sich daher nicht im Stande, die Fragen zu beantworten. Die ausgefüllten Fragebogen stammen daher vermutlich von Richtern, die sich tatsächlich mit ähnlichen Fällen in der Praxis befassen.
- 161 Die zweite Umfrage im Herbst 2004 erfolgte nach derselben Methode, allerdings gab es hier kein Vorbild für den Fragebogen (es gibt natürlich Vorbilder für die Fragestellungen; es handelt sich dabei durchwegs um Fragen, die in der psychologischen Literatur schon wiederholt gestellt und die auf juristische Sachverhalte angepasst wurden). Es wurden bis auf eine Ausnahme (Rückschaufehler) andere kognitive Täuschungen als bei der ersten Umfrage geprüft (der Rückschaufehler wurde mit einem neuen Sachverhalt untersucht). Wiederum gab es von jedem Sachverhalt mindestens zwei bis maximal vier Varianten (zweifaktorieller Versuchsplan bei einzelnen Fragen). Es wurden diesmal 476 Fragebogen an die Richterinnen und Richter der Straf- und Zivilgerichte der Kantone Basel-Land-

³⁴² Der Begleitbrief war von Prof. Dr. Manfred Rehbinder unterschrieben, so dass niemand ahnen konnte, wer der eigentliche Urheber der Umfrage war.

³⁴³ ANDRÉ KUHN/PATRICE VILLET/ALINE JOYET/FLORIAN WILLI, *Öffentliche Meinung und Strenge der Richter*, Crimscope Nr. 19, Lausanne 2000, 2, berichten von einer Rücklaufquote von 44 %. Sie legten 290 Richterinnen und Richtern vier Strafrechtsfälle zur Beurteilung vor.

³⁴⁴ RAINER SCHNELL/PAUL B. HILL/ELKE ESSER, *Methoden der empirischen Sozialforschung*, 7. Aufl. München 2005, 306 ff.

schaft, Basel-Stadt, Bern und Graubünden (mit Ausnahme des italienischsprachigen Bezirksgerichts Moesa) verschickt. Titelblatt und Begleitbrief waren weitgehend identisch mit der ersten Umfrage. Anders als beim ersten Mal wurde diesmal nach einer Woche ein persönlich adressiertes Erinnerungsschreiben versandt. Da die Umfrage anonym erfolgte und ich daher nicht wissen konnte, welche Richter den Fragebogen retourniert hatten, haben alle Richter die Erinnerung erhalten; dies ist einzelnen sauer aufgestossen, die den Fragebogen bereits retourniert hatten, liess sich aber nicht vermeiden. 234 Fragebogen wurden zumindest teilweise ausgefüllt zurückgeschickt, was einer Rücklaufquote von 49 % entspricht; die Erinnerung scheint daher zumindest eine gewisse Wirkung gehabt zu haben.

II. Auswertung, oder: was bedeutet „statistisch signifikant“?

- 162 Im Folgenden wird ein beobachteter Unterschied gemäss üblicher Konvention in den Sozialwissenschaften dann als signifikant bezeichnet, wenn der p-Wert kleiner als 0,05 ist.³⁴⁵ Ein p-Wert von weniger als 0,05 wurde auch schon – in durchaus kritischer Absicht – als der „heilige Gral des Hypothesentestens“ bezeichnet.³⁴⁶ Was bedeutet der p-Wert?
- 163 Als erstes muss man klar unterscheiden zwischen „statistisch signifikant“ und dem umgangssprachlichen „wesentlich“ oder „erheblich“. Das Ergebnis eines Versuchs kann statistisch signifikant, aber unwesentlich sein.³⁴⁷ Wenn die Anzahl der untersuchten Fälle sehr gross ist, sind bereits kleine Differenzen statistisch signifikant, aber möglicherweise praktisch bedeutungslos. In dieser Arbeit wird der Ausdruck „signifikant“ ausschliesslich für Ergebnisse verwendet, die im Sinne der herkömmlichen Definition statistisch signifikant sind; hingegen wird ein Effekt als „wesentlich“ oder „erheblich“ beschrieben, wenn er – nach der subjektiven Meinung des Autors – praktisch bedeutsam ist.
- 164 Angenommen, es gibt a priori zwei Hypothesen, um die beobachteten Daten zu erklären. In der Untersuchung zum Darstellungseffekt (S. 95 ff.) beurteilten die Richter beispielsweise den gleichen Sachverhalt, aber eine Gruppe der Richter beurteilte ihn aus Sicht der Klägerin, eine zweite Gruppe aus Sicht der Beklagten. In der Gruppe der Richter, die den Fall aus der Sicht der Klägerin beurteilten, empfahlen 56,6 % (43 von 76), einen Vergleich abzuschliessen. Nur 43 % der Richter (43 von 100), die den Fall aus der Perspektive der Beklagten beurteilten, rieten jedoch der Beklagten zum Vergleich. Die Arbeitshypothese ist, dass die Perspektive einen Einfluss auf die Vergleichsbereitschaft (genauer gesagt die Risikoaversion) hat. Die Nullhypothese ist, dass es keinen Unterschied in der Vergleichsbereitschaft der beiden Gruppen gibt und dass der beobachtete Unterschied auf Zufall zurückzuführen ist.
- 165 Man kann, soviel ist sicher, zwei Fehler machen: die Nullhypothese verwerfen, obwohl sie wahr ist (Fehler 1. Art, auch α -Fehler), oder die Nullhypothese beibehalten, obwohl die

³⁴⁵ SCHNELL/HILL/ESSER, FN 344, 450.

³⁴⁶ PAULA BRAITSTEIN, Give P's a Chance: Understanding p-values in scientific research, Living+ Jan/Feb 2004, 34 (erhältlich unter www.bcpwa.org/articles/issue_28_34_give_ps_a_chance.pdf; besucht am 1. März 2005).

³⁴⁷ SCHNELL/HILL/ESSER, FN 344, 452.

Arbeitshypothese zutrifft (Fehler 2. Art, auch β -Fehler).³⁴⁸ Im Strafverfahren wäre ein Fehler 1. Art, einen Unschuldigen zu verurteilen, während ein Fehler 2. Art darin bestünde, einen Schuldigen freizusprechen.³⁴⁹ Mit diesem Beispiel wird auch klar, dass die Konsequenzen der beiden Fehler durchaus nicht dieselben sein müssen – die meisten Menschen sind sich einig, dass es schwerer wiegt, einen Unschuldigen einzusperren, als einen Verbrecher laufen zu lassen.

- 166 Wie wahrscheinlich ist es nun, dass die im Experiment zum Darstellungseffekt beobachteten Daten auf Zufall zurückzuführen sind? Zum Verständnis kann man folgendes Gedankenexperiment machen:³⁵⁰ angenommen, es gäbe keinen Unterschied der Perspektive, sondern es gäbe einfach Richter, die von ihrer Persönlichkeit her eher zum Vergleich raten und Richter, die eher den Prozess empfehlen. Mit anderen Worten hätten 86 der 176 Richter, die die Frage beantwortet haben, ohnehin einen Vergleich empfohlen, ganz unabhängig davon, ob sie den Fall aus der Perspektive der Klägerin oder der Beklagten beurteilten. Die Wahrscheinlichkeit, dass *alle* 86 Richter, die zum Vergleich rieten, sich *zufällig* in derselben Gruppe befinden, ist offensichtlich sehr klein: Es ist nicht relevant, welcher Gruppe der erste vergleichsfreudige Richter angehört. Der zweite vergleichsfreudige Richter muss aber in der gleichen Gruppe wie der erste Richter sein, die Wahrscheinlichkeit dafür ist 0,5. Der dritte Richter muss wiederum in der gleichen Gruppe sein wie die ersten beiden Richter, die Wahrscheinlichkeit dafür beträgt ebenfalls 0,5. Da die beiden Ereignisse unabhängig sind, ergibt sich die Wahrscheinlichkeit, dass der zweite *und* dritte Richter in der gleichen Gruppe wie der erste Richter sind, aus der Produktregel: $0,5 \cdot 0,5 = 0,25$. Bei 86 vergleichsfreudigen Richtern, die alle in der gleichen Gruppe sein müssen, beträgt die Wahrscheinlichkeit daher $0,5^{86}$.³⁵¹
- 167 Die Wahrscheinlichkeit, dass man alle 86 vergleichsbereiten Richtern in einer Gruppe findet, wenn die Nullhypothese zutrifft und die Verteilung rein zufällig ist, beträgt daher $0,5^{86}$ (eine Null, gefolgt von einem Komma und 25 Nullen, ehe eine 2 folgt). Diese Wahrscheinlichkeit entspricht dem p-Wert, dieser ist nämlich definiert als die Wahrscheinlichkeit, dass sich die beobachteten Daten so wie beobachtet oder extremer realisieren, wenn die Nullhypothese zutrifft.³⁵² Anders formuliert gibt der p-Wert die Wahrscheinlichkeit an, einen Fehler 1. Art zu begehen.
- 168 Offensichtlich wird die Wahrscheinlichkeit, dass sich die Daten wie beobachtet oder extremer realisieren, grösser, je weniger extrem die beobachteten Daten sind. Im Experiment zum Darstellungseffekt haben 43 Richter in der Gruppe „Klägerin“ und 43 in der Gruppe „Beklagte“ zum Vergleich geraten, was auf den ersten Blick nicht gerade auf einen statistisch signifikanten Einfluss der Gruppenzugehörigkeit hindeutet. Allerdings war die

³⁴⁸ LUDWIG FAHRMEIER/RITA KÜNSTLER/IRIS PIGEOT/GERHARD KUNZ, Statistik, 5. Aufl. Berlin etc. 2004, 416. Bekannt ist auch der Fehler 3. Art – die beiden Fehlerarten zu verwechseln.

³⁴⁹ Unter der Annahme, dass „unschuldig“ die Nullhypothese ist, die im ersten Fall (zu Unrecht) verworfen wird.

³⁵⁰ Die Darstellung folgt HANS ZEISEL/DAVID KAYE, Prove it with Figures: Empirical Methods in Law and Litigation, New York 1997, 83 f.

³⁵¹ Da eine Gruppe nur 76 Mitglieder hatte, beträgt die Wahrscheinlichkeit, dass alle Mitglieder dieser Gruppe vergleichsfreudige Richter sind, „nur“ $0,5^{76}$. Die kleinere Gruppe schwächt den Signifikanztest.

³⁵² FAHRMEIER/KÜNSTLER/PIGEOT/KUNZ, FN 348, 420.

Gruppe der Richter, die den Fall aus Sicht der Beklagten beurteilten, grösser (100 verglichen mit 76 in der Gruppe „Klägerin“). Intuitiv ergibt es Sinn, dass die Verteilung in diesem Fall nicht mehr rein zufällig ist – der Anteil der vergleichsbereiten Richter macht in der einen Gruppe 56,6 %, in der anderen 43 % aus; zu erwarten wären 48 % in jeder Gruppe. Wie sehr die Verteilung von der unter der Nullhypothese erwarteten gleichmässigen Verteilung abweicht, kann (bei kategorialen Daten wie im Beispiel) der χ^2 -Test (Chi-Quadrat-Test) beantworten.³⁵³ In diesem Fall beträgt der χ^2 Wert (bei einem Freiheitsgrad) 3,187. Dies entspricht einem p-Wert von 0,074.³⁵⁴ Mit anderen Worten würde man, wenn man den gleichen Versuch 100 Mal wiederholen würde, in etwas mehr als sieben Fällen Daten beobachten, die gleich oder extremer als die beobachteten Daten sind. Der „heilige Gral“ der statistischen Signifikanz wird daher in diesem Fall knapp verfehlt.

- 169 Aus einem Signifikanztest, der zu einem Ergebnis gelangt, das „auf dem 5 % Niveau“³⁵⁵ signifikant ist, darf man unter keinen Umständen schliessen, dass die Arbeitshypothese mit einer Wahrscheinlichkeit von 95 % zutrifft (oder die Nullhypothese mit 95 %-iger Wahrscheinlichkeit falsch ist). Der p-Wert gibt die Wahrscheinlichkeit an, dass die Daten so extrem oder extremer als beobachtet sind *unter der Annahme, dass die Nullhypothese wahr ist*; formal $P(\text{extremere Daten} \mid \text{Nullhypothese})$.³⁵⁶ Die bedingte Wahrscheinlichkeit, dass die Nullhypothese falsch ist, ist $P(\text{Nullhypothese} \mid \text{extremere Daten})$, und diese Wahrscheinlichkeit ist *nicht* $1 - P(\text{extremere Daten} \mid \text{Nullhypothese})$. Dieser Fehler, im strafrechtlichen Zusammenhang als „Trugschluss des Anklägers“ bezeichnet, ist nicht selten (mehr dazu hinten, S. 135 ff.).
- 170 Ob die Nullhypothese falsch ist, hängt nicht nur von den beobachteten Daten ab, sondern auch davon, wie wahrscheinlich die Nullhypothese *a priori* war, d.h. bevor die beobachteten Daten bekannt wurden. Ob die Verwerfung der Nullhypothese bedeutet, dass die Arbeitshypothese (mit hoher Wahrscheinlichkeit) zutrifft, hängt weiterhin davon ab, ob nur die Arbeitshypothese die beobachteten Daten erklären kann oder auch eine andere – möglicherweise nicht bedachte – Hypothese. Wie wahrscheinlich die Nullhypothese *a priori* ist und welche anderen Hypothesen die beobachteten Daten erklären können, lässt sich – anders als der p-Wert – nicht mathematisch berechnen. Das Testen von Hypothesen unter Berücksichtigung der Anfangswahrscheinlichkeit und alternativer Hypothesen (nach dem Bayes-Theorem, näheres hinten, S. 125 ff.) führt daher ein subjektives Element in die anscheinend so objektive Statistik ein. Dies mag mit ein Grund sein, warum in den meisten sozialwissenschaftlichen Veröffentlichungen nur der p-Wert angegeben wird.
- 171 Neben dem χ^2 Test für kategoriale Daten werden in dieser Arbeit der t-Test (für normalverteilte Daten) und der Wilcoxon-Rangsummen-Test (resp. der äquivalente Mann-

³⁵³ FAHRMEIER/KÜNSTLER/PIGEOT/KUNZ, FN 348, 445 ff.

³⁵⁴ Wie ein Blick in die entsprechende Tabelle in einem Statistik-Lehrbuch (z.B. FAHRMEIER/KÜNSTLER/PIGEOT/KUNZ, FN 348, 583 ff.), oder, heutzutage einfacher, das Computerprogramm zeigt.

³⁵⁵ Der p-Wert, ab dem von „Signifikanz“ gesprochen wird, bezeichnet man als „Signifikanzniveau“, daher die Formulierung. Dass dieses Niveau bei 5 % liegt, ist eine mehr oder weniger arbiträre Konvention, an die man sich in den Sozialwissenschaften weitgehend hält.

³⁵⁶ ZEISEL/KAYE, FN 350, 81.

Whitney-U-Test) für ordinale, nicht parametrische Daten verwendet.³⁵⁷ Alle Berechnungen wurden mit SPSS 11.0 für Windows durchgeführt.

- 172 Wo dies möglich ist (für nicht nominale Daten) werden das arithmetische Mittel, der Median und die Standardabweichung angegeben. Das arithmetische Mittel – die Summe der Werte geteilt durch die Anzahl der Werte – wird dem üblichen Sprachgebrauch gemäss als Durchschnitt oder Schnitt bezeichnet. Der Median ist der Wert, der in der Mitte der Reihe liegt, wenn man eine Reihe von Messwerten der Größe nach sortiert.³⁵⁸ Die eine Hälfte der Werte ist größer, die andere Hälfte kleiner als der Median. Im Gegensatz zum Durchschnitt verändert sich der Median durch einzelne Extremwerte kaum; er ist daher geeigneter als der Durchschnitt, die Verteilung der Daten zu beschreiben, wenn einzelne Extremwerte den Durchschnitt verzerren (dies ist beispielsweise bei den Resultaten zum Ankereffekt der Fall, S. 83 ff.). Die Standardabweichung schliesslich ist ein Mass für die Streuung der Daten um das arithmetische Mittel. Je grösser die Standardabweichung, desto weiter weg sind die einzelnen Werte im Schnitt vom arithmetischen Mittel. Da für die Berechnung der Standardabweichung die Abweichungen vom Mittelwert quadriert werden, steigt die Standardabweichung bei einzelnen Ausreissern schnell stark an.³⁵⁹

III. Interne Validität

- 173 Eine Studie ist intern valide, wenn die manipulierte Variable (das Treatment oder die unabhängige Variable) tatsächlich für die Varianz der abhängigen Variablen verantwortlich ist. Werden die Messwerte hingegen durch einen oder mehrere Störfaktoren verändert, dann ist die interne Validität, oder Gültigkeit, verletzt. Sind sowohl Störfaktoren als auch der Stimulus für die beobachteten Effekte verantwortlich, spricht man von einer „Konfundierung“ der Effekte.³⁶⁰
- 174 Wenn der einzige (systematische) Unterschied zwischen Versuchs- und Kontrollgruppe im Treatment besteht, ist die Annahme zulässig, dass ein beobachteter Unterschied der abhängigen Variablen zwischen Versuchs- und Kontrollgruppe auf das Treatment zurückzuführen ist. Durch Zufallszuweisung (Randomisierung) von Untersuchungseinheiten in Versuchs- und Kontrollgruppe kann garantiert werden, dass die Unterschiede zwischen Versuchs- und Kontrollgruppe rein zufällig sind.³⁶¹ Bei der vorliegenden Studie wurden die Richterinnen und Richter zufällig der Versuchs- oder Kontrollgruppe zugewiesen. Zwar wurden keine „echten“ Zufallszahlen verwendet, um die Zuweisung vorzunehmen, aber die Fragebogen wurden gemischt und dann zusammen mit dem Begleitbrief in Couverts verpackt. Es gibt keinen Grund anzunehmen, dass sich die Richter, die eine Version der Frage erhalten haben, systematisch von den Richtern, die eine andere Version erhalten haben, unterscheiden. Wenn ein Unterschied in der abhängigen Variablen besteht – wenn z.B. die

³⁵⁷ Mehr zu diesen Tests bei FAHRMEIER/KÜNSTLER/PIGEOT/KUNZ, FN 348, 437 ff. (t-Test) und 459 ff. (Wilcoxon-Rangsummen-Test).

³⁵⁸ FAHRMEIER/KÜNSTLER/PIGEOT/KUNZ, FN 348, 55.

³⁵⁹ Zu Berechnung der Standardabweichung siehe FAHRMEIER/KÜNSTLER/PIGEOT/KUNZ, FN 348, 69.

³⁶⁰ SCHNELL/HILL/ESSER, FN 344, 219.

³⁶¹ SCHNELL/HILL/ESSER, FN 344, 223.

Richter, bei denen der Kläger eine hohe Genugtuungssumme beantragt, systematisch mehr zusprechen als die Richter, denen kein bezifferter Antrag vorliegt – darf man daher annehmen, dass der Unterschied zwischen den beiden Gruppen auf das Treatment, in diesem Fall den Antrag des Klägers, zurückzuführen ist. M. a.W. gibt es keinen vernünftigen Grund, anzunehmen, dass die vorliegende Studie nicht intern valide ist.

IV. Externe Validität

- 175 Problematischer als die interne ist die externe (oder ökologische) Validität der vorliegenden Studie. Diese ist definiert als Möglichkeit der Generalisierung der experimentellen Resultate auf andere Personen(-gruppen) und Situationen.³⁶² Dass Richter eine höhere Genugtuungssumme zusprechen, wenn sie nach der Lektüre eines Sachverhalts von einer halben Seite mit einem Antrag des Klägers von Fr. 3 Mio. konfrontiert werden, ist zwar interessant. Eigentlich interessiert aber natürlich, ob dieser Effekt auch in einem Gerichtsverfahren wirkt, in dem sehr viel mehr und andere Informationen zur Verfügung stehen als in einem kurzen schriftlichen Sachverhalt.
- 176 Einige Richterinnen und Richter bezweifeln, dass es möglich ist, aufgrund der Studie Aussagen über reale juristische Entscheidungen zu machen. „Ich möchte Ihre Fachkompetenz nicht anzweifeln, habe aber Mühe mit der Vorstellung, dass Antworten auf diese fünf Fragen gültige Aussagen in Bezug auf juristische Entscheidungen in der Realität geben können!“ schrieb eine Richterin. Andere schickten den Fragebogen unausgefüllt zurück mit der Bemerkung, dass „die Fragen weit an den Entscheidungssituationen vorbeigehen, welche mir als Richter begegnen“ oder sie nicht in der Lage seien, aufgrund der spärlichen Unterlagen ein einigermaßen seriöses Urteil zu fällen und daher auch nicht glaubten, mit dem Ausfüllen des Fragebogens einen Beitrag zur Erforschung der richterlichen Entscheidungsfindung leisten zu können. Diese kritische Haltung gegenüber sozialwissenschaftlicher Forschung ist nicht untypisch für Richter.³⁶³
- 177 Kritik an rechtspsychologischer Forschung mit kurzen schriftlichen Fallskizzen (so genannten „Vignetten“) wurde wiederholt geäußert und ist ernst zu nehmen. Am pointier-testen haben wohl KONECNI/EBBESEN die Kritik vorgebracht.³⁶⁴
- 178 Die Gründe, die gegen die externe Validität solcher Studien sprechen, sind wohl bekannt. Die Entscheidungen, die von den Versuchspersonen in Simulationsstudien getroffen werden, haben keine Konsequenzen. Die Fallskizzen reduzieren die Art und den Umfang der zur Verfügung stehenden Informationen in einem Ausmass, die als Karikatur echter Gerichtsfälle – die sich eher durch ein Übermass an Informationen auszeichnen – erscheint. Häufig werden die Stimuli zudem in „zerlegter“ Form dargeboten, um einzelne Faktoren manipulieren zu können. Die Analyse eines komplexen Sachverhalts ist aber in der Praxis häufig der schwierigste Teil der juristischen Entscheidungsfindung. Oft fehlen in Experi-

³⁶² SCHNELL/HILL/ESSER, FN 344, 219.

³⁶³ RICHARD E. REDDING/N. DICKSON REPPUCH, Effects of Lawyers' Socio-political Attitudes on Their Judgments of Social Science in Legal Decision Making, Law and Human Behavior 1999, 31-54, 47.

³⁶⁴ VLADIMIR J. KONECNI/EBBE B. EBBESEN, Social Psychology and the Law: The Choice of Research Problems, Settings, and Methodology, in: KONECNI/EBBESEN (Hrsg.), FN 265, 27-44.

menten prozedurale Schritte wie die Beratung in der Gruppe, die einen Einfluss auf das Urteil haben können. Schliesslich entsprechen auch die abhängigen Variablen – beispielsweise die Beurteilung der Schuld auf einer numerischen Skala – nicht den Entscheidungen (schuldig ja/nein), die in Wirklichkeit zu treffen sind.³⁶⁵

- 179 KONECNI/EBBESEN gehen davon aus, dass der Urteilende eine aufgabenspezifische (*task-specific*) Entscheidungsregel für jede konkrete Entscheidung konstruiert.³⁶⁶ Gemäss dieser Auffassung kann es keine Theorie geben, die a priori erklärt, wie ein Urteilender eine bestimmte Aufgabe anpacken wird. KONECNI/EBBESEN stellen daher nicht nur die externe Validität von „*paper and pencil*“ Studien in Frage, sondern auch diejenige realistischerer Studien, die zum Beispiel Videoaufzeichnungen der Aussagen von Parteienanwälten und Zeugen verwenden und Beratung vor der Entscheidung erlauben.³⁶⁷ Sie untermauern diese Aussage empirisch mit drei eigenen Simulationsstudien, die anschliessend mit den Resultaten von Feldstudien verglichen wurden.³⁶⁸ In keinem Fall, so EBBESEN/KONECNI, hätten die Resultate der Simulationsstudien erlaubt, die Resultate der Feldstudien richtig vorauszusagen.³⁶⁹ Sie kommen daher zum Schluss, dass es „gefährlich und nahezu unverantwortlich [ist], auf der Grundlage von Simulationsstudien mit von ihrem lebenswirklichen Kontext losgelösten Effekten Schlussfolgerungen und Empfehlungen für das Rechtssystem zu formulieren“.³⁷⁰
- 180 Für KONECNI/EBBESEN können nur Feldstudien, d.h. Untersuchungen in der Lebenswirklichkeit, zu validen Resultaten kommen. Dabei ziehen sie Archivstudien der Beobachtung von Gerichtsverhandlungen aus verschiedenen methodischen Gründen vor,³⁷¹ wobei sie empfehlen, dass das Rechtssystem differenzierte Verfahren zur Erhebung von Daten zum eigenen Handeln am besten gleich einbaut.³⁷² Die Schlussfolgerungen und Empfehlungen von KONECNI/EBBESEN werden von Vertretern theorielastiger Soziologie oder Psychologie begrüsst; so z.B. von LÖSCHPER in ihrer diskursanalytischen Habilitationsschrift „Bausteine für eine Theorie richterlichen Urteilens“, die Simulationsstudien als „Experimente im sozialen Vakuum“ kritisiert.³⁷³
- 181 Trotz der Kritik von KONECNI/EBBESEN und anderen, bleiben Simulationsstudien unvermindert populär in der rechtspsychologischen Forschung. Häufig werden die beson-

³⁶⁵ KONECNI/EBBESEN, FN 364, 28.

³⁶⁶ VLADIMIR J. KONECNI/EBBE B. EBBESEN, A Critique of Theory and Method in Social-Psychological Approaches to Legal Issues, in: BRUCE DENNIS SALES (Hrsg.), *The Trial Process*, New York 1981, 481-498, 488.

³⁶⁷ KONECNI/EBBESEN, FN 364, 31.

³⁶⁸ EBBE B. EBBESEN/VLADIMIR J. KONECNI, On the External Validity of Decision-Making Research: What Do We Know About Decisions in the Real World? in: THOMAS S. WALLSTEN (Hrsg.), *Cognitive Processes in Choice and Decision Behavior*, Hillsdale 1980, 21-45.

³⁶⁹ EBBESEN/KONECNI, FN 368, 25 ff.

³⁷⁰ VLADIMIR J. KONECNI/EBBE B. EBBESEN, Methodische Probleme in der Forschung über juristische Entscheidungsprozesse – unter besonderer Berücksichtigung experimenteller Simulationen, *Gruppendynamik* 1991, 175-188, 179.

³⁷¹ KONECNI/EBBESEN, FN 366, 493; KONECNI/EBBESEN, FN 364, 35 f.

³⁷² KONECNI/EBBESEN, FN 370, 185.

³⁷³ GABRIELE LÖSCHPER, *Bausteine für eine psychologische Theorie richterlichen Urteilens*, Baden-Baden 1999, 20.

ders unrealistischen „*paper and pencil*“ Studien mit Studierenden als Versuchspersonen durchgeführt. KERR und BRAY haben 1982 die Methoden von 72 Simulationsstudien mit Geschworenen (*mock juries*) untersucht und festgestellt, dass rund 50 % der Studien mit kurzen schriftlichen Fallskizzen und studentischen „Geschworenen“ durchgeführt wurden.³⁷⁴ Der Anteil von Simulationsstudien an den in *Law and Human Behavior* zwischen 1977 und 1996 publizierten Studien zum Verhalten von Geschworenen ist effektiv gestiegen, von rund 50 % auf 70-80 %. Der Anteil der Simulationsstudien mit Studierenden und kurzen Fallskizzen ist ebenfalls gestiegen, zu Lasten aufwändigerer Studien mit Videoaufnahmen und Bürgern, die als Geschworene qualifiziert sind (*jury eligible citizens*).³⁷⁵ Der Trend geht also genau in die entgegengesetzte als die von den Kritikern geforderte Richtung.

- 182 Es gibt zynische, methodologische und pragmatische Gründe für die ungebrochene Popularität von Simulationsstudien. Der zynische Grund ist, dass nur akademische Karriere macht, wer viel publiziert. Studien mit Sachverhaltsskizzen und studentischen Versuchspersonen sind viel schneller und billiger als aufwändige Simulations- oder gar Feldstudien, die in der gleichen Anzahl Publikationen resultieren.³⁷⁶ Solange die Anzahl der Publikationen ein ausschlaggebendes Kriterium für akademische Berufungen ist, werden einfache Simulationsstudien beliebt bleiben.
- 183 Der methodologische Grund für Simulationsstudien liegt darin, dass die experimentelle Kontrolle, die der Versuchsleiter bei Simulationen ausüben kann, grosse Vorteile hat. Sie führt, wie oben dargelegt, in der Regel zu einer hohen internen Validität. Ein Einfluss konfundierender Variablen kann meist ausgeschlossen werden.³⁷⁷ Wollte man beispielsweise den Einfluss des klägerischen Antrags auf die zugesprochene Genugtuungssumme im Feld studieren, so müsste man alle objektiven Faktoren, die die Höhe der Genugtuung beeinflussen (wie Schwere der Verletzung des Opfers und Schwere des Verschuldens des Täters) konstant halten. Tut man dies nicht – wie beispielsweise die beiden in Fn. 422 und 425 erwähnten spanischen Studien – so ist eine beobachtete Korrelation von Antrag des Klägers und zugesprochener Genugtuungssumme vermutlich nicht auf den Ankereffekt, sondern auf die Schwere der Verletzung zurückzuführen – wer schwerer verletzt ist, verlangt mehr und erhält, zu Recht, mehr zugesprochen. Es gibt statistische Techniken, den Einfluss möglicherweise konfundierender Variablen konstant zu halten, z.B. die multiple Regressionsanalyse. Diese setzt aber eine relativ hohe Anzahl Fälle voraus, um zu valablen Resultaten zu gelangen. Da die Zuspreehung von Genugtuung in der Schweiz nicht sonderlich häufig ist, müsste man die Rechtsprechung zahlreicher Kantone während mehrerer Jahre auswerten, um überhaupt genügend Daten für eine valide statistische Auswertung zu erhalten.

³⁷⁴ ROBERT M. BRAY/NORBERT L. KERR, Methodological Considerations in the Study of the Psychology of the Courtroom, in: NORBERT L. KERR/ROBERT M. BRAY (Hrsg.), *The Psychology of the Courtroom*, San Diego etc. 1982, 287-323, 294.

³⁷⁵ BRIAN H. BORNSTEIN, The Ecological Validity of Jury Simulations: Is the Jury Still Out?, *Law and Human Behavior* 1999, 75-92, 87.

³⁷⁶ KONECNI/EBBESEN, FN 366, 486.

³⁷⁷ BRAY/KERR, FN 374, 296; AMIRAM ELWORK/BRUCE DENNIS SALES/DAVID SUGGS, The Trial: A Research Review, in: SALES (Hrsg.), FN 366, 1-59, 51.

- 184 Ein weiterer Vorteil von Simulationsstudien ist, dass sie – anders als Feldstudien – repliziert werden können.³⁷⁸ Während eine einzelne Studie nie genügt, um externe Validität zu generieren, kann die systematische und variierte Replikation einer Studie dazu beitragen, dass man darauf vertrauen darf, dass die Resultate auch in Situationen, die nicht konkret untersucht wurden, Gültigkeit haben.³⁷⁹ Viele der in der vorliegenden Studie gestellten Fragen wurden beispielsweise schon Studierenden gestellt. Zeigen zu können, dass auch Richterinnen und Richter teilweise die gleichen Fehler machen wie studentische Versuchspersonen, ist ein weiterer Puzzlestein, der die externe Validität der den Fragen zu Grunde liegenden psychologischen Theorie stützt. Dabei besteht ein offensichtlicher Trade-off zwischen Realitätsnähe des Stimulus-Materials (also Vignetten statt den gesamten Akten eines Prozesses oder der Videoaufzeichnung mündlicher Plädoyers und Zeugenaussagen) und der Realitätsnähe der Versuchspersonen. Wer sich dafür entscheidet, mit Richtern als Versuchspersonen zu arbeiten, muss sich fast zwangsweise auf kurze Sachverhalte beschränken. Richter sind nicht bereit, sich mehrere Stunden in einen Sachverhalt einzulesen, um an einer sozialwissenschaftlichen Studie teilzunehmen. Die Vermutung sei erlaubt, dass die gleichen Richter, die an der vorliegenden Studie nicht teilgenommen haben, weil sie zu unrealistisch sei, an einer realistischeren Studie nicht teilgenommen hätten, weil sie den dafür notwendige Zeitaufwand als unzumutbar erachten.
- 185 Den Vorteilen der Simulationsstudien stehen Nachteile der Feldstudien gegenüber. Erstens sind gerade die von KONECNI/EBBESSEN favorisierten Archivstudien wegen datenschutzrechtlicher Bestimmungen in Europa häufig schwierig oder gar nicht durchzuführen. Auch werden zahlreiche interessierende Faktoren durch die gerichtlichen Akten nicht erhoben – beispielsweise wird in der Schweiz das letzte Vergleichsangebot der Parteien vor dem Prozess nicht erfasst, so dass es ausgeschlossen ist, das letzte Vergleichsangebot mit der im Urteil zugesprochenen Summe zu vergleichen, wie dies RACHLINSKI getan hat (mehr dazu hinten, S. 92 ff.).³⁸⁰ Die Fallzahlen sind, wie bereits erwähnt, oft zu gering, um quantitative Aussagen machen zu können. Selbst in Deutschland beschränken sich Feldstudien auf häufige Delikte wie den einfachen Diebstahl, da andere Fälle nicht oft genug vorkommen.³⁸¹ In der Schweiz ist dieses Problem noch gravierender als in Deutschland. So ist es beispielsweise ausgeschlossen, mittels einer Feldstudie quantitative Aussagen über den Einfluss einer Gesetzesänderung auf die Anzahl der jährlich ausgesprochenen Verwahrungen zu machen (dazu S. 243 ff.), wenn jährlich im Schnitt nur rund 20 stationäre Massnahmen und Verwahrungen ausgesprochen werden.³⁸² Schliesslich stellen die schriftlichen Zeugnisse, die das Justizsystem produziert, keinen repräsentativen Querschnitt der tatsäch-

³⁷⁸ BRAY/KERR, FN 374, 296.

³⁷⁹ BRAY/KERR, FN 374, 301.

³⁸⁰ RACHLINSKI, FN 219.

³⁸¹ OSWALD, FN 260, 15.

³⁸² Bundesamt für Statistik, Strafurteilsstatistik, Stand der Datenbank 12. August 2004, Verurteilungen nach Art der Massnahme (der Durchschnitt der Verwahrungen nach Art. 42 StGB [Gewohnheitsverbrecher] und der stationären Massnahmen und Verwahrungen nach Art. 41 Ziff. 1 [psychisch abnorme Täter] der letzten fünf Jahre [1999-2003, neuste verfügbare Zahlen] betrug 22,6; die Statistik unterscheidet bei den Massnahmen nach Art. 43 Ziff. 1 leider nicht zwischen stationären Massnahmen und Verwahrungen).

lichen Fälle dar.³⁸³ Die – faktisch überaus zahlreichen – Fälle, die verglichen werden, hinterlassen entweder gar keine (bei aussergerichtlichem Vergleich) oder nur geringe (bei gerichtlichem Vergleich) Spuren in den Gerichtsarchiven, aus denen sich kaum wertvolle Erkenntnisse gewinnen lassen. Wenn man sich daher für den Einfluss psychologischer Faktoren auf die Vergleichsempfehlung durch Richter interessiert (S. 93 ff.), kann man sich nicht auf Archivstudien verlassen.

- 186 Der pragmatische Grund für Simulationsstudien sind die prohibitiv hohen Kosten von Feldstudien. Die konsequente Umsetzung der Forderung von KONECNI/EBBESSEN würde dazu führen, dass kaum mehr rechtspsychologische Forschung betrieben würde. Forschungsgelder für gross angelegte Feldstudien sind kaum vorhanden. Wie BRAY/KERR bemerken: „Wir werden nicht viele Erkenntnisse gewinnen, wenn wir jede Forschung verbieten ausser derjenigen, die prohibitiv teuer ist für die meisten Forschenden“.³⁸⁴ Eine Simulationsstudie mit möglicherweise eingeschränkter externer Validität ist dieser Auffassung nach, der ich mich anschliessen möchte, gar keiner Studie immer noch vorzuziehen.³⁸⁵ Selbst KONECNI/EBBESSEN können auf Simulationsstudien nicht verzichten: rund ein Drittel der empirischen Studien, die in der von ihnen herausgegebenen Aufsatzsammlung „*The Criminal Justice System*“ erschienen sind, sind Simulationsstudien.
- 187 Bisher wurde dargelegt, warum Simulationsstudien *trotz* ihrer eingeschränkten Validität einer Feldstudie unter Umständen vorzuziehen sind. Wie aber steht es denn nun um die externe Validität von Simulationsstudien? Die Frage, die sich nur empirisch beantworten lässt, ist nur ungenügend erforscht. KONECNI/EBBESSEN kommen wie erwähnt zu einem eher pessimistischen Resultat. Andere Forscher haben nicht die Resultate von Feldstudien mit derjenigen von Simulationsstudien verglichen, sondern die Resultate unterschiedlich realitätsnaher Simulationen. Die Wahl der Versuchspersonen, also Studierende (*college undergraduates*) statt als Geschworene wählbare Bürger (*jury eligible citizens*), hat nach den übereinstimmenden Resultaten von mehreren Dutzend Studien keinen Einfluss auf die beobachteten Effekte.³⁸⁶ Erstaunlicherweise hat auch die Verwendung von Videotapes statt kurzer Sachverhaltsskizzen als Stimulus-Material keinen wesentlichen Einfluss auf die Resultate; allerdings ist die Anzahl der Studien, die dies untersuchen, geringer und die Ergebnisse nicht so eindeutig wie bei den Versuchspersonen.³⁸⁷ Wer wie KONECNI/EBBESSEN jede Form von Simulation ablehnt, da Entscheidungsregeln aufgabenspezifisch und situativ seien, wird sich dadurch natürlich nicht überzeugen lassen. Wer diesen Standpunkt allerdings ernsthaft vertritt, muss auf jede generelle Theorie zur Erklärung menschlichen Entscheidungsverhaltens unter Unsicherheit, und damit auf jegliche Generalisierbarkeit, verzichten. Für diejenigen, die nicht so weit gehen wollen, ist die Tatsache, dass sich die Ergebnisse von Studien mit kurzen Sachverhaltsskizzen auf aufwändigere

³⁸³ SHARI SEIDMAN DIAMOND, The Challenges of Socio-Legal Research on Decision Making: Psychological Successes and Failures, *Journal of Law and Society* 1995, 78-84, 79 f.

³⁸⁴ BRAY/KERR, FN 374, 316: „We do not stand to gain much knowledge by prohibiting all research except that which is restrictingly expensive for most investigators“.

³⁸⁵ Ebenso ELWORK/SALES/SUGGS, FN 377, 1-59, 51

³⁸⁶ BORNSTEIN, FN 375, 78.

³⁸⁷ BORNSTEIN, FN 375, 82.

Studien mit Videoaufnahmen generalisieren lassen, zumindest ein Indiz dafür, dass es um die externe Validität von „*paper and pencil*“ Studien nicht so schlecht stehen kann.

- 188 Im vorliegenden Fall kommt hinzu, dass sich die Fragestellungen auf ein psychologisches Forschungsprogramm stützen, das seit dreissig Jahren mit einer enormen Flut von empirischen Studien erforscht wird. Obwohl die Randbedingungen der einzelnen Urteilsheuristiken weiterhin umstritten sind, handelt es sich bei den hier besprochenen Phänomenen doch um Effekte, die allgemein als robust gelten.³⁸⁸ Insofern die Resultate der Studie bestimmte Denkweisen aufzeigen, ist es wahrscheinlich, dass die Richter auch bei der Beurteilung tatsächlicher Fälle gleich denken, zumal sie sich der Effekte meist nicht bewusst sein dürften. Erhöhte Motivation allein ist kein Allheilmittel. Nur wenn die erhöhte Aufmerksamkeit und die längere Überlegungszeit dazu führen, dass die kognitive Täuschung erkannt und vermieden wird, verbessert sich die Entscheidungsqualität. Dies darf nicht leichthin angenommen werden.³⁸⁹ Der beste Beweis dafür sind Studien, in denen kognitive Illusionen selbst dann die Entscheidungen negativ beeinflussen, wenn es für die Probanden um sehr viel Geld geht. Der Darstellungseffekt beispielsweise wirkt sogar, wenn die (chinesischen) Testpersonen um Einsätze in der Höhe von zwei Monatslöhnen spielen.³⁹⁰
- 189 Kognitive Täuschungen wurden auch in Feldstudien ausserhalb des Labors und/oder in kontrollierten Experimenten mit sehr realistischem Stimulus-Material nachgewiesen.³⁹¹ Kognitive Täuschungen beeinflussen auch das Urteil von Experten ausserhalb des Labors.³⁹² Die Häufigkeit des Vorkommens einer Krankheit wird von Ärzten auch dann nicht genügend beachtet, wenn sie in der Klinik Patienten diagnostizieren.³⁹³ MUSSWEILER und Kollegen schickten einen Schauspieler mit einem echten Opel Kadett bei 60 (uneingeübten) Automechanikern und –händlern vorbei, um seinen Wert schätzen zu lassen. Die Schätzung wurde erheblich durch den vom Schauspieler manipulierten Ankereffekt beeinflusst.³⁹⁴ Sehr viel realitätsnäher kann eine Studie nicht mehr sein. Diese Generalisierbarkeit der Resultate der Forschung zu den *heuristics and biases* auf viele natürliche Entscheidungssituationen legt es nahe anzunehmen, dass sie sich auch auf Juristen und juristische Entscheidungen generalisieren lassen. Natürlich kann man die Auffassung vertreten, dass Richter anders als alle anderen sind und den Täuschungen, denen fast alle Experten unterliegen, nicht unterliegen, oder dass spezifische Elemente des gerichtlichen Verfahrens verhindern, dass sich kognitive Täuschungen auf juristische Urteile auswirken. Wer diese Behauptung aber aufstellt, trägt in Anbetracht der empirischen Forschung der letzten dreissig

³⁸⁸ Detaillierte Nachweise bei den einzelnen Heuristiken.

³⁸⁹ AMOS TVERSKY/DANIEL KAHNEMAN, FN 325, 274.

³⁹⁰ KACHELMEIER/SHEHATA, FN 304.

³⁹¹ Siehe generell die im Kapitel “Real-World Applications” abgedruckten Studien in: GILOVICH/GRIFFIN/KAHNEMAN, 601-762; CAMERER, FN 303.

³⁹² RACHLINSKI, FN 456, 154; DEREK J. KOEHLER/LYLE BRENNER/DALE GRIFFIN, The Calibration of Expert Judgment: Heuristics and Biases Beyond the Laboratory, in: GILOVICH/GRIFFIN/KAHNEMAN (Hrsg.), 686-715.

³⁹³ DAWES, FN 616, 425.

³⁹⁴ THOMAS MUSSWEILER/FRITZ STRACK/TIM PFEIFFER, Overcoming the Inevitable Anchoring Effect: Considering the Opposite Compensates for Selective Accessibility, Personality and Social Psychology Bulletin 2000, 1142-1150, 1146

sig Jahre die Beweislast dafür.³⁹⁵ Mit der Behauptung alleine ist es nicht getan, solange ihr jegliche empirische Basis fehlt.

- 190 Zusammengefasst gesagt sind Feldstudien Simulationsstudien dort vorzuziehen, wo sie aus ökonomischen Gründen durchführbar und entsprechende Daten vorhanden sind. In den häufigen Fällen, in denen dies nicht der Fall ist, sind Simulationsstudien eine *second best* Lösung mit eigenen Vorteilen, namentlich der grösseren internen Validität von echten Experimenten. Die Resultate aufwändiger Simulationsstudien unterscheiden sich dabei kaum von simplen Studien, die kurze Sachverhaltsskizzen als Stimulus-Material verwenden. Kognitive Täuschungen wurden in einer Reihe von Studien ausserhalb des Labors unter natürlichen Bedingungen und bei Experten beobachtet, so dass man davon ausgehen darf, dass es sich um robuste, nicht nur unter Laborbedingungen auftretende Phänomene handelt. Wer die Auffassung vertritt, dass sie ausgerechnet im Justizsystem keine Rolle spielen, muss in Anbetracht der Erkenntnisse aus dreissig Jahren Forschung den empirischen Nachweis dafür erbringen.

³⁹⁵ Zur „Beweislastumkehr“ beim Vorliegen empirischer rechtssoziologischer Daten MANFRED REHBINDER, Rechtssoziologie, 5. Aufl. München 2003, Rz. 24 und 59.

§ 8 Ankereffekt (*anchoring*)

- 191 Als Ankereffekt (*anchoring*) wird die Assimilation eines numerischen Urteils an einen vorgegebenen Vergleichsstandard bezeichnet. Der Vergleichsstandard wirkt wie ein „Anker“ für die endgültige Schätzung und zieht diese in seine Richtung.³⁹⁶
- 192 In einer inzwischen berühmten frühen Studie zum Ankereffekt baten AMOS TVERSKY und DANIEL KAHNEMAN die Teilnehmenden zu schätzen, wie viel Prozent der Mitgliedstaaten der Vereinten Nationen afrikanische Länder sind.³⁹⁷ Zuerst mussten die Teilnehmenden beantworten, ob die wirkliche Prozentzahl über oder unter einer Zahl lag, die durch das Drehen eines „Glücksrades“ bestimmt wurde. Das Glücksrad war – ohne dass die Partizipanten dies wussten – so manipuliert worden, dass es entweder bei 10 oder bei 65 stoppte. Natürlich antworteten alle Teilnehmenden, bei denen das Rad bei 65 stoppte, dass der Prozentsatz afrikanischer Länder in der UNO unter 65% liegen würde. Aber, dies die nächste Frage, welches ist der *exakte* Wert?
- 193 Es zeigte sich, dass das (angeblich) zufällige Ergebnis des Glücksrads, obwohl es offensichtlich keinen Hinweis auf die richtige Antwort geben konnte, die Antworten massgeblich beeinflusste: Wenn das Rad bei 10 stoppte, lag die Schätzung der Versuchspersonen für den prozentualen Anteil der afrikanischen Länder in der UNO im Median bei 25 %; stoppte das Rad bei 65, lag die Schätzung im Median bei 45 %.³⁹⁸
- 194 Selbst absurd hohe Anker können das Urteil beeinflussen. In einer Studie von GEORGE QUATTRONE et al. schätzten Studentinnen den Wert eines Schulbuches höher ein, wenn sie zuerst gefragt wurden, ob der Wert höher oder tiefer als \$ 7'128.53 sei.³⁹⁹ Ebenso schätzten sie die durchschnittliche Jahrestemperatur in San Francisco höher, wenn sie zuerst entscheiden mussten, ob sie unter oder über 558° Fahrenheit (rund 292° Celsius) liegt.⁴⁰⁰
- 195 Der Ankereffekt ist ausserordentlich robust. Aufklärung über den Ankereffekt und die ausdrückliche Warnung, sich nicht von einem vorgängig genannten Wert beeinflussen lassen, zeigen keine Wirkung.⁴⁰¹ Der Ankereffekt wirkt unbeabsichtigt und unbewusst. Finanzielle

³⁹⁶ TVERSKY/ KAHNEMAN, FN 8, 1128 ff.; PLOUS, *The Psychology of Judgment and Decision Making*, New York 1993, 145 ff.; HELMUT JUNGERMANN/HANS-RÜDIGER PFISTER/KATRIN FISCHER, *Die Psychologie der Entscheidung*, Heidelberg 1998, 171 f.

³⁹⁷ TVERSKY/KAHNEMANN, FN 8, 1128.

³⁹⁸ TVERSKY/KAHNEMANN, FN 8, 1128. Die richtige Antwort ist übrigens für 1973, als die Studie durchgeführt wurde, 35 %.

³⁹⁹ G.A. QUATTRONE/C.P. LAWRENCE/D.L. WARREN/K. SOUZA-SILVA/S.E. FINKEL/D.E. ANDRUS, *Explorations in Anchoring: The Effects of Prior Range, Anchor Extremity, and Suggestive Hints*, unpubl. Manuskript, Stanford University, 1984 (rapportiert in REID HASTIE/ROBYN M. DAWES, *Rational Choice in an Uncertain World: The Psychology of Judgement and Decision Making*, Thousand Oaks, 2001, 103).

⁴⁰⁰ HASTIE/DAWES, FN 399, 103.

⁴⁰¹ TIMOTHY D. WILSON/CHRISTOPHER E. HOUSTON/KATHRYN M. ETLING/NANCY BREKKE, *A New Look at Anchoring Effects: Basic Anchoring and Its Antecedents*, *Journal of Experimental Psychology: General* 1996, 387-402, 397.

Belohnungen für eine möglichst genaue Antwort haben daher keinen Einfluss auf die Stärke des Ankereffekts.⁴⁰²

- 196 Fachwissen schützt ebenfalls nicht vor dem Ankereffekt. In der bereits erwähnten Studie von MUSSWEILER und Kollegen teilte der Schauspieler der Hälfte der Automechaniker und -händler mit, ein Freund habe den Wert des Wagens auf DM 5'000 (hoher Anker) resp. DM 2'800 (tiefer Anker) geschätzt. Die Experten, die mit dem hohen Anker konfrontiert waren, schätzten den Wert des Autos im Median auf DM 3'347, diejenigen, die mit dem tiefen Anker konfrontiert waren, auf DM 2'652.⁴⁰³ NORTHCROFT/NEALE berichten von einem Experiment, in dem eine Gruppe von Immobilienagenten gebeten wurde, den Wert eines Hauses zu schätzen. Die Agenten erhielten alle Informationen, die gemeinhin für eine Immobilienschätzung benötigt werden, wie Lage und Merkmale des Hauses und die Preise, die für vergleichbare Liegenschaften in der näheren Umgebung bezahlt worden waren, und hatten die Gelegenheit, das Haus zu inspizieren. Die Agenten in der Gruppe mit dem hohen Listenpreis (= Anker) schätzten das Haus im Schnitt \$ 7'000 (rund 10 %) höher ein als die Kontrollgruppe.⁴⁰⁴
- 197 Erfahrung hat schliesslich ebenfalls keinen mässigenden Einfluss auf den Ankereffekt. Die Versuchspersonen in einem Experiment von ILANA RITOV, die über Preis und Lieferbedingungen für eine Warenlieferung verhandelten, liessen sich auch nach mehrmaliger Wiederholung unvermindert stark vom ersten Angebot beeinflussen.⁴⁰⁵
- 198 Auch Zeitablauf hat nur einen geringen Einfluss auf die Stärke des Ankereffekts.⁴⁰⁶ MUSSWEILER fragte die Hälfte seiner Versuchspersonen in einem klassischen Ankereffekt-Experiment, ob die Durchschnittstemperatur in Deutschland über oder unter 5° C; die andere Hälfte, ob sie unter oder über 20° C liege. Eine Woche nach dieser Frage gaben die Versuchspersonen, die mit dem hohen Anker konfrontiert waren, eine Schätzung für den tatsächlichen Wert von 16,44° C ab, die Versuchspersonen, die mit dem tiefen Anker konfrontiert waren, eine solche von 9,91° C.⁴⁰⁷
- 199 Einen mässigenden Einfluss auf den Ankereffekt hat das Wissen der schätzenden Person über den tatsächlichen Zielwert. Das leuchtet unmittelbar ein, wenn man sich den Extremfall vorstellt, in dem die Person den richtigen Wert mit Sicherheit kennt: kein Richter wird den Streitwert für die Berufung an das Bundesgericht in vermögensrechtlichen Zivilstreit-

⁴⁰² WILLIAM WRIGHT/URTON ANDERSON, Effects of Situation Familiarity and Financial Incentives on Use of the Anchoring and Adjustment Heuristic for Probability Assessment, *Organizational Behavior and Human Decision Processes* 1989, 68-82; WILSON/HOUSTON/ETLING/BREKKE, FN 401, 396; TVERSKY/KAHNEMAN, FN 8, 1128.

⁴⁰³ MUSSWEILER/STRACK/PFEIFFER, FN 394, 1146.

⁴⁰⁴ GREGORY B. NORTHCROFT/MARGARET A. NEALE, Experts, Amateurs, and Real Estate: An Anchoring-and-Adjustment Perspective on Property Pricing Decisions, *Organizational Behavior & Human Decision Processes* 1987, 84-97, 95 f.

⁴⁰⁵ ILANA RITOV, Anchoring in Simulated Competitive Market Negotiation, *Organizational Behavior and Human Decision Processes* 1996, 16-25.

⁴⁰⁶ Spätestens dann, wenn sich die Versuchspersonen nicht mehr an den Anker erinnern können, verliert er natürlich seinen Einfluss.

⁴⁰⁷ THOMAS MUSSWEILER, The Durability of Anchoring Effects, *European Journal of Social Psychology* 2001, 431-442, 438.

tigkeiten höher schätzen, wenn man ihn vorher fragt, ob der wahre Wert über oder unter Fr. 100'000 liegt. Die eindeutige Antwort auf die Frage steht im Gesetz (Art. 46 Bundesgesetz über die Organisation der Bundesrechtspflege) und dürfte den meisten Richtern bekannt sein; eine Beeinflussung durch den Anker ist daher ausgeschlossen. MUSSWEILER und STRACK zeigen, dass sich Menschen umso weniger vom Anker beeinflussen lassen, je mehr sie über den Zielwert wissen, selbst wenn sie den Zielwert nicht mit Sicherheit kennen.⁴⁰⁸

- 200 Der Ausgangswert bewirkt, dass man sich ernsthaft überlegt, ob der korrekte Wert in der Nähe des Ankers liegen könnte und sich Umstände vorstellt, die dies als möglich erscheinen lassen, während man Umstände, die gegen den Ankerwert sprechen, übersieht.⁴⁰⁹ Werden die Versuchspersonen aufgefordert, Gründe aufzuführen, die gegen den Ankerwert sprechen, verringert sich der Einfluss des Ankers, verschwindet aber erstaunlicherweise selbst dann nicht vollständig.⁴¹⁰ Wer mit einem unmöglich hohen Anker konfrontiert wird, überlegt sich, wie hoch der tatsächliche Wert im Extremfall sein könnte. Gründe, die für diesen hohen, aber denkbaren, Wert sprechen, werden gesucht. Die Antwort tendiert dann zum hohen, aber gerade noch vorstellbaren, Wert hin.⁴¹¹

I. Ankereffekt vor Gericht

- 201 In Gerichtsverfahren gibt es Anker in der Form von Rechtsbegehren, Anträgen des Staatsanwalts und Angeboten zur vergleichweisen Erledigung des Streites. Es liegt daher einigermassen nahe, den Einfluss des Ankereffekts auf den Ausgang von Gerichtsverfahren zu untersuchen.
- 202 SAKS/KIDD (1980) und FITZMAURICE/PEASE (1986) waren ersichtlich die ersten, die sich Gedanken zum Einfluss des Ankereffekts auf richterliche Entscheidungen machten. SAKS/KIDD wiesen darauf hin, dass schlaue Anwälte den Ankereffekt zu Gunsten ihrer Klienten einsetzen können.⁴¹² FITZMAURICE/PEASE befürchten, dass die Urteile der Berufungsgerichte als Anker für die erstinstanzlichen Gerichte wirken und dies dazu führt, dass immer härtere Strafen ausgesprochen werden. Sie argumentieren, dass nur die extremsten Fälle, in denen das erstinstanzliche Gericht den Strafraum voll ausgeschöpft und möglicherweise überschritten hat, an die Berufungsinstanz gelangen.⁴¹³ Die Entscheide des *Court of Appeal*, die sich mit dem Strafmass befassen, betreffen daher in der Regel Extremfälle. Da der *Court of Appeal* in Ermessenfragen das erstinstanzliche Gericht nur zurückhaltend korrigieren wird, liegen seine Urteile in Strafsachen daher an der äussersten Grenze des Zulässigen. Wenn sich die unteren Gerichte bewusst oder unbewusst an diesen

⁴⁰⁸ THOMAS MUSSWEILER/FRITZ STRACK, Numeric Judgments under Uncertainty: The Role of Knowledge in Anchoring, *Journal of Experimental Social Psychology* 2000, 495-518.

⁴⁰⁹ THOMAS MUSSWEILER/FRITZ STRACK, Explaining the Enigmatic Anchoring Effect: Mechanisms of Selective Accessibility, *Journal of Personality and Social Psychology* 1997, 437-446, 437 f.

⁴¹⁰ MUSSWEILER/STRACK/PFEIFFER, FN 403, 1148 f.

⁴¹¹ THOMAS MUSSWEILER/FRITZ STRACK, Considering the Impossible: Explaining the Effects of Implausible Anchors, *Social Cognition* 2001, 145-160.

⁴¹² SAKS/KIDD, FN 183, 141.

⁴¹³ Ob diese Annahme auch in der Schweiz zutrifft, bezweifle ich.

Besonderer Teil

Werten orientieren, besteht die Gefahr, dass die ausgesprochenen Strafen immer mehr eskalieren und wiederum das Empfinden, was eine angemessene Strafe ist, beeinflussen. FITZMAURICE/PEASE befürchten eine „*folie à deux*“, die zu immer härteren Strafen führt und ihres Erachtens nur dadurch verhindert wird, dass sich die erstinstanzlichen Gerichte wenig an den Urteilen des *Court of Appeal* orientieren.⁴¹⁴

- 203 Weder SAKS/KIDD noch FITZMAURICE/PEASE untermauern ihre Gedanken zum Einfluss des Ankereffekts mit empirischen Studien. Zwischenzeitlich sind jedoch zahlreiche empirische Untersuchungen über den Einfluss des Ankereffekts auf Gerichtsverfahren publiziert worden. Diese Studien zeigen, dass Anker einen erheblichen Einfluss auf den Ausgang des Verfahrens haben können.
- 204 Zahlreiche Studien zeigen einen Einfluss der Höhe des klägerischen Begehrens auf die durch die (von Studierenden gespielten) Geschworenen zugesprochene Schadenersatzsumme.⁴¹⁵ In einer Studie sprachen die Geschworenen dem Kläger durchschnittlich \$ 188'462 zu, wenn sein Anwalt \$ 300'000 verlangt hatte, jedoch – bei ansonsten gleichem Sachverhalt – durchschnittlich \$ 421'000, wenn der Anwalt \$ 700'000 verlangt hatte.⁴¹⁶ Kläger, die sehr viel verlangen, werden zwar als unsympathisch empfunden – sie erhalten aber trotzdem mehr zugesprochen als die sympathischen, bescheidenen Kläger.⁴¹⁷
- 205 Auch Richter sind nicht gegen den Einfluss des Ankereffekts gefeit. Eine deutsche Studie von BIRTE ENGLISH und THOMAS MUSSWEILER und die bereits erwähnte amerikanische Untersuchung von GUTHRIE/RACHLINSKI/WISTRICH zeigen, dass auch sie dem Ankereffekt unterliegen.⁴¹⁸
- 206 Die *Magistrate Court Judges* in der Studie von GUTHRIE et al. sprachen dem Kläger im Median \$ 1 Mio. Schadenersatz zu, wenn kein Antrag gestellt wurde. Behauptete der Beklagte, die Streitwertgrenze von \$ 75'000 für die sachliche Zuständigkeit sei nicht er-

⁴¹⁴ FITZMAURICE/PEASE, FN 185, 21 f.

⁴¹⁵ ALLAN RAITZ/EDITH GREENE/JANE GOODMAN/ELIZABETH F. LOFTUS, Determining Damages: The Influence of Expert Testimony on Jurors' Decision Making, *Law and Human Behavior* 1990, 385-395; VERLIN B. HINSZ/KRISTIN E. INDAHL, Assimilation to Anchors for Damage Awards in a Mock Civil Trial, *Journal of Applied Social Psychology* 1995, 991-1026; GRETCHEN B. CHAPMAN/BRIAN H. BORNSTEIN, The More You Ask For, the More You Get: Anchoring in Personal Injury Verdicts, *Applied Cognitive Psychology* 1996, 519-540; JENNIFER K. ROBBENNOLT/CHRISTINA A. STUDEBAKER, Anchoring in the Courtroom: The Effects of Caps on Punitive Damages, *Law and Human Behavior* 1999, 353-373; REID HASTIE/DAVID A. SCHADKE/JOHN W. PAYNE, Juror Judgements in Civil Cases: Effects of Plaintiff's Requests and Plaintiff's Identity on Punitive Damage Awards, *Law and Human Behavior* 1999, 445-469; JOHN MALOUFF/NICOLA S. SCHUTTE, Shaping Juror Attitudes: Effects of Requesting Different Damage Amounts in Personal Injury Trials, *Journal of Social Psychology* 491-497.

⁴¹⁶ MALOUFF/SCHUTTE, FN 415, 495.

⁴¹⁷ CHAPMAN/BORNSTEIN, FN 415, 526 f.

⁴¹⁸ BIRTE ENGLISH/THOMAS MUSSWEILER, Sentencing under uncertainty: Anchoring effects in the Court Room, *Journal of Applied Social Psychology* 2001, 1535-1551; GUTHRIE/RACHLINSKI/WISTRICH, FN 20, 792; siehe auch EBBE B. EBBESEN/VLADIMIR J. KONECNI, Decision Making and Information Integration in the Courts: The Setting of Bail, *Journal of Personality and Social Psychology* 1975, 805-821, die zeigen, dass die Höhe der Kaution in erster Linie vom Antrag des Bezirksanwalts abhängt.

reicht und auf die Klage sei nicht einzutreten, wiesen zwar alle Richter diesen Antrag ab. Sie sprachen dann aber, beeinflusst vom tiefen Anker, im Median nur \$ 882'000 zu.⁴¹⁹

- 207 ENGLISH/MUSSWEILER baten deutsche Strafrichterinnen mit durchschnittlich 15 Jahren Berufserfahrung, aufgrund einer vierseitigen Schilderung einer Vergewaltigung⁴²⁰ das Strafmass für den Täter festzulegen. Verlangte der „Ankläger“ – gemäss den Unterlagen ein Informatikstudent im ersten Semester – 34 Monate Freiheitsstrafe, so lag das Strafmass im Median bei 35,75 Monaten. Verlangte der Ankläger zwölf Monate, wurde der Täter im Schnitt zu 28 Monaten verurteilt.⁴²¹ Obwohl der Ankläger offensichtlich keine Ahnung von Strafrecht hatte, liessen sich die erfahrenen Richter von seinem Antrag beeinflussen.
- 208 Zwei Feldstudien aus Spanien kommen zum Schluss, dass der Ankereffekt auch tatsächliche Urteile zu beeinflussen vermag. EUGENIO GARRIDO MARTIN und CARMEN HERRERO ALONSO berichten, dass die Urteile spanischer Richter bei „*estupro*“ (entspricht ungefähr Sex mit minderjährigen Abhängigen) und Vergewaltigung mit dem Antrag des Staatsanwalts hoch korrelieren.⁴²² Sie kommen zum Schluss, dass dies die Hypothese, dass der Ankereffekt die Gerichtsurteile beeinflusst, stützt.⁴²³ Bei anderen Sexualdelikten wie versuchter Vergewaltigung und Sex mit Minderjährigen ist die Übereinstimmung weniger gross. GARRIDO/HERRERO schreiben dies dem Druck der Öffentlichkeit zu, die die ersten beiden Delikte für besonders verabscheuenswürdig hält. Der Richter würde sich in diesen Fällen nicht getrauen, vom Antrag des Staatsanwalts abzuweichen, da er den Zorn der Öffentlichkeit fürchten müsse.⁴²⁴
- 209 Die Studie von GARRIDO/HERRERO ist zweifellos interessant, scheint mir aber wegen methodischer Defizite nicht geeignet, den Einfluss des Ankereffekts zu beweisen. Die hohe Korrelation zwischen dem Antrag des Staatsanwalts und dem richterlichen Urteil kann auch auf andere Faktoren als auf den Ankereffekt zurückzuführen sein. Korrelation bedeutet bekanntlich nicht Kausation. Die Ursache der hohen Übereinstimmung kann z.B. darin liegen, dass Staatsanwalt und Richter unabhängig voneinander den Fall ähnlich beurteilen. Ohne dass man Variablen wie Schwere der Tat, Anzahl der Vorstrafen des Täters, Alter des Opfers etc. kontrolliert, kann man aus einer Korrelation von Strafantrag und Urteil nichts über die Ursache dieser Korrelation ableiten.
- 210 Eine weitere Feldstudie aus Spanien von FARIÑA et al. kommt aufgrund der Analyse von 555 Strafurteilen zum Schluss, dass 63,6 % davon vom Ankereffekt beeinflusst sind.⁴²⁵

⁴¹⁹ GUTHRIE/RACHLINSKI/WISTRICH, FN 20, 792.

⁴²⁰ Die Schilderung enthielt Stellungnahmen einer Ärztin, eines Psychologen, zweier Zeugen, des Angeklagten und des Opfers, siehe ENGLISH/ MUSSWEILER, FN 418, 1539.

⁴²¹ ENGLISH/MUSSWEILER, FN 418, 1545.

⁴²² EUGENIO GARRIDO MARTIN/CARMEN HERRERO ALONSO, Influence of the Prosecutor's Plea on the Judge's Sentencing in Sexual Crimes: Hypothesis of the Theory of Anchoring by Tversky and Kahneman, in SANTIAGO REDONDO/VICENTE GARRIDO/JORGE PÉREZ/ROSEMARY BARBERET, Advances in Psychology and Law, Berlin etc. 1997, 215-226, 221.

⁴²³ GARRIDO/HERRERO, FN 422, 223.

⁴²⁴ GARRIDO/HERRERO, a. a. O.

⁴²⁵ FRANCISCA FARIÑA/RAMON ARCE/MERCEDES NOVO, Anchoring in Judicial Decision-Making, Psychology in Spain 2003, 56-65.

FARIÑA und Kollegen erfassen neben dem Antrag des Staatsanwalts noch zahlreiche weitere Variablen, die in erster Linie die Begründungsdichte des Urteils messen.⁴²⁶ Sie können zeigen, dass die Begründungsdichte abnimmt, je mehr sich das Urteil dem Strafantrag der Anklage annähert.⁴²⁷ Sie schliessen daraus, dass der Ankereffekt seine Aufgabe als Urteilsheuristik – als einfache Faustregel, die komplexe Urteile vereinfacht – wie vorgesehen erfüllt. Dort, wo die Gerichte aufgrund der Faustregel urteilen, können sie sich lange Begründungen sparen.⁴²⁸

- 211 Auch diese Studie ist interessant, leidet aber m.E. an den gleichen methodischen Defiziten wie die Studie von GARRIDO/HERRERO. Dass die Begründungsdichte von Urteilen, in denen eine grosse Übereinstimmung zwischen Antrag und Urteil besteht, geringer ist, lässt sich dadurch erklären, dass dies die einfachen Fälle sind, bei denen es über die Schwere der Schuld des Angeklagten wenig Zweifel gibt. In diesen Fällen gibt es weder Grund, vom Antrag des Staatsanwalts abzuweichen, noch Grund, das Urteil besonders sorgfältig zu begründen. Da Variablen, die die Schwere der Tat messen, auch in dieser Studie nicht kontrolliert wurden, scheint mir der Schluss, dass „alle Anzeichen dafür sprechen, dass Richter die Ankerheuristik verwenden und objektivere Formen der Informationsverarbeitung unterdrücken“, gewagt.⁴²⁹ Ich zweifle zwar nicht daran, dass der Ankereffekt auch auf tatsächliche Urteile einen Einfluss hat, aber die beiden spanischen Studien vermögen diese Annahme nicht zu belegen.

II. Eigene Studie zum Ankereffekt

A. Frage zum Ankereffekt

- 212 Den Richterinnen und Richtern der Kantone Aargau, St. Gallen und Zürich wurde folgender Sachverhalt zur Beurteilung vorgelegt (2003):

Das Auto der Klägerin wurde von einem Lastwagen erfasst. Der Lastwagen hatte der Klägerin den Rechtsvortritt verweigert. Der Lastwagenfahrer war im Zeitpunkt des Unfalles betrunken. Seine Blutalkoholkonzentration betrug 1,6 Promille. Die Klägerin, eine 20-jährige Frau, fuhr korrekt und konnte den Unfall nicht vermeiden.

Die Klägerin, die vor dem Unfall in guter körperlicher Verfassung war, ist durch die unmittelbaren Folgen des Unfalles querschnittgelähmt. Eine Heilung ist ausgeschlossen. Nach rund dreimonatigem Aufenthalt im Spital konnte sie nach Hause entlassen werden. Sie wird ihren erlernten Beruf als Detailhandelsverkäuferin nicht mehr ausüben können und wird auch auf ihr Hobby, das Reiten, verzichten müssen.

Die Klägerin hat sich mit der Haftpflichtversicherung des Halters des Lastwagens aussergerichtlich über die Höhe des Schadenersatzes geeinigt. Die beklagte Versicherung anerkennt auch den Anspruch der Klägerin auf Genugtuung im Grundsatz. Streitig ist zu diesem Zeitpunkt einzig noch die Höhe der Genugtuung.

⁴²⁶ FARIÑA/ARCE/NOVO, FN 425, 58.

⁴²⁷ FARIÑA/ARCE/NOVO, FN 425, 60.

⁴²⁸ FARIÑA/ARCE/NOVO, FN 425, 62.

⁴²⁹ FARIÑA/ARCE/NOVO, FN 425, 61: „In brief, all the indications are that judicial decision-makers employ the anchoring heuristic to subordinate more objective forms of information processing”.

- 213 Dieser Sachverhalt enthält in verkürzter Form Angaben zu allen gemäss bundesgerichtlicher Rechtsprechung massgeblichen Faktoren der Genugtuungsbemessung. Gemäss BGE 112 II 131, 133 kommt es „vor allem auf die Art und Schwere der Verletzung, die Intensität und die Dauer der Auswirkungen auf die Persönlichkeit des Betroffenen, sowie auf den Grad des Verschuldens an, das den Schädiger am Unfallereignis trifft“.
- 214 In der ersten Version des Fragebogens stellen weder Klägerin noch Beklagte einen Antrag (Gruppe „kein Anker“, N = 65). In der zweiten Version des Fragebogens stellt die Klägerin den Antrag, es sei ihr eine Genugtuungssumme von Fr. 3 Mio. zuzusprechen (Gruppe „hoher Anker“, N = 45). Diese Forderung ist in Anbetracht der zurückhaltenden schweizerischen Praxis klar überrissen.⁴³⁰ In der dritten Version des Fragebogens schliesslich stellt die Klägerin einen Antrag von Fr. 3 Mio, während die Beklagte Fr. 10'000 anbietet (Gruppe „hoher/tiefer Anker“, N = 62). Das Angebot der Beklagten wiederum ist in Anbetracht der Praxis klar zu tief. Bei allen drei Versionen des Fragebogens waren Sachverhalt und zu beantwortende Frage identisch. Die Richter wurden gefragt, wie hoch ihrer Meinung nach eine angemessene Genugtuung in diesem Fall sei, soweit sie dies aufgrund der Angaben im Sachverhalt beurteilen könnten.
- 215 Die Hypothese ist natürlich, dass sich die Richterinnen und Richter durch den (unplausibel) hohen Anker beeinflussen lassen. Die dritte Version der Frage diene dazu, zu testen, ob der tiefe Antrag der Beklagten den Ankereffekt des hohen klägerischen Antrags kompensieren würde. Träfe dies zu, so würde der Ankereffekt zumindest in kontradiktorischen Verfahren – und somit in den meisten Zivil- und Strafprozessen – keine wesentliche Rolle spielen.

B. Resultate

- 216 Der Antrag der Klägerin zeigte einen sehr starken Einfluss auf die zugesprochene Genugtuungssumme. Der Median der Genugtuung lag bei der Gruppe ohne Anker bei Fr. 100'000, bei der Gruppe mit dem hohen Anker bei Fr. 400'000 und bei der Gruppe mit einem hohen und tiefen Anker bei Fr. 200'000 (siehe Tabelle 7). Die Verteilung der zugesprochenen Summen war rechtsschief (*positively skewed*), wie Abbildung 8 unschwer erkennen lässt.

Gruppe	1. Quartil (25% Quartil)	2. Quartil (Median)	3. Quartil (75% Quartil)	Standard- abweichung
Kein Anker (N = 65)	Fr. 50'000	Fr. 100'000	Fr. 150'000	Fr. 308'880
Hoher Anker (N = 45)	Fr. 150'000	Fr. 400'000	Fr. 1'000'000	Fr. 902'704
Hoher/tiefer Anker (N = 62)	Fr. 100'000	Fr. 200'000	Fr. 687'500	Fr. 760'174

Tabelle 7: Resultate Ankereffekt

- 217 Drei Viertel der Richter, die mit dem hohen Antrag der Klägerin konfrontiert waren, sprachen eine höhere Genugtuungssumme zu als drei Viertel der Richter, die in der Gruppe

⁴³⁰ Siehe die bei KLAUS HÜTTE/PETRA DUCKSCH, Die Genugtuung, Zürich 2003, Tabelle VIII, wiedergegebenen Entscheide, die darauf schliessen lassen, dass die Genugtuung im erwähnten Fall bei rund Fr. 100'000 bis 150'000 liegen würde; eine genauere Angabe ist naturgemäss nicht möglich.

ohne Anker waren. Der Antrag der Beklagten hatte zwar einen mässigenden Einfluss auf die zugesprochenen Genugtuungssummen, vermochte diese aber bei weitem nicht auf das Niveau der ohne Anker zugesprochenen Genugtuungen zu reduzieren – der Median lag immer noch um das Doppelte über der in der Referenzgruppe ohne Anker im Median zugesprochenen Summe.

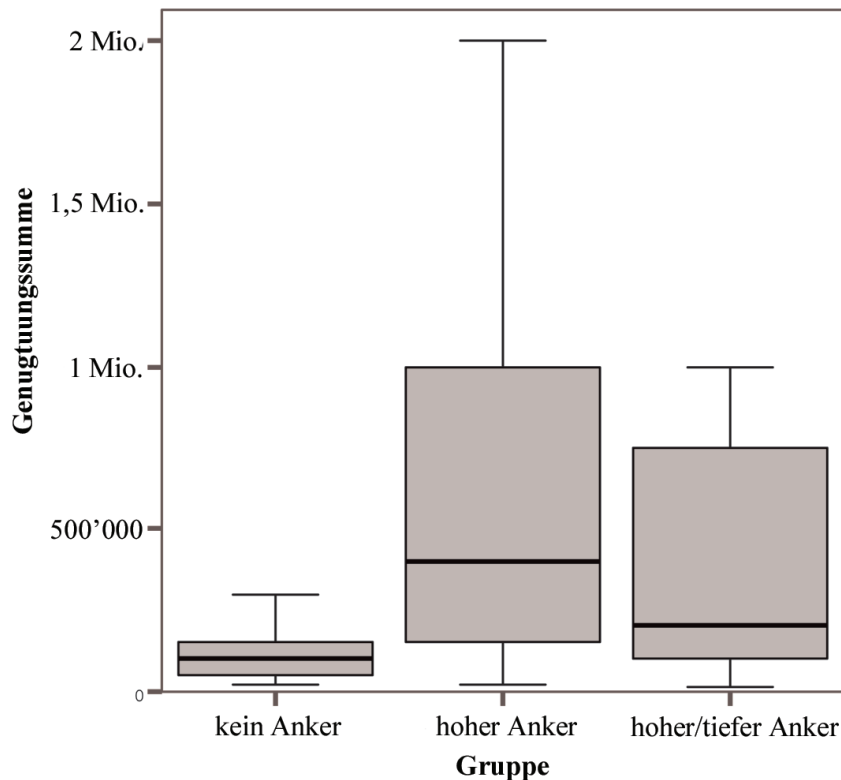


Abbildung 8: Resultate Ankereffekt

- 218 Der Unterschied zwischen der Gruppe ohne Anker und der Gruppe mit hohem Anker⁴³¹ sowie zwischen der Gruppe ohne Anker und der Gruppe mit hohem/tiefem Anker ist statistisch signifikant.⁴³² Der Unterschied zwischen der Gruppe mit hohem Anker und der Gruppe mit hohem/tiefen Anker ist hingegen statistisch nicht signifikant.⁴³³ Dies deutet darauf hin, dass der „zweite“ Anker den Wert nicht signifikant zu beeinflussen vermag.
- 219 Von den neun Richtern, die den Fragebogen zurückschickten, aber diese Frage nicht beantworteten, bemerkten drei, dass sich die Genugtuungssumme aufgrund der Angaben im Sachverhalt nicht beurteilen liesse. Ein Richter meinte, er könne ohne Konsultation der Präjudizien keine sinnvolle Antwort geben.

⁴³¹ Mann-Whitney U = 679,5; Wilcoxon W = 2'824,5; p < 0.001.

⁴³² Mann-Whitney U = 1'078,5; Wilcoxon W = 3'225,5; p < 0.001.

⁴³³ Mann-Whitney U = 1'203,5; Wilcoxon W = 3'156,5; p = 0.225.

C. Diskussion

- 220 Der Antrag der Klägerin vermochte die Beurteilung der Richter erheblich zu beeinflussen. Der Antrag mag dazu geführt haben, dass die Richter sich überlegten, ob besondere Umstände für einen hohen Genugtuungsanspruch sprachen und die Lücken in der Sachverhaltsdarstellung entsprechend ergänzten. Diese Erklärung steht in Übereinstimmung mit der psychologischen Literatur zum Ankereffekt.

1. Wie das kontradiktorische Verfahren den Einfluss des Ankereffekts verringern kann

- 221 Inwieweit lassen sich die Resultate der Umfrage auf die Wirklichkeit übertragen? Sicherlich führt der überrissene Antrag der Klägerin in einem wirklichen Gerichtsverfahren nicht dazu, dass sie in einem Viertel der Fälle eine Genugtuung von mehr als Fr. 687'500 zugesprochen bekommt. Es gibt sowohl psychologische wie rechtliche Faktoren, die dem entgegenwirken.
- 222 Die grössere Detailliertheit der Sachverhaltsdarstellung in einem Gerichtsverfahren führt dazu, dass weniger Raum für mentale Ergänzungen des Sachverhalts bleibt, die den Anker rechtfertigen. Immerhin zeigen aber die Studien von MUSSWEILER/ STRACK/PFEIFFER und NORTHCROFT/NEALE, dass der Ankereffekt auch im wirklichen Leben wirkt, wenn die Experten über alle nur denkbaren Informationen verfügen und das zu schätzende Objekt in Natura begutachten können.⁴³⁴ Es ist daher unwahrscheinlich, dass die zusätzlichen Informationen, die in einem Gerichtsverfahren generiert werden, den Ankereffekt gänzlich verschwinden lassen. Auch der Zeitablauf zwischen Verhandlung und Urteil verhindert den signifikanten Einfluss des Ankers nicht,⁴³⁵ ebenso wenig wie die Erfahrung der Richterinnen und Richter.⁴³⁶
- 223 Einen grösseren praktischen Einfluss in einem kontradiktorischen Verfahren hat sicher die Generierung von gegen den ersten Antrag sprechenden Argumenten. MUSSWEILER/STRACK/PFEIFFER zeigen, dass sich der Ankereffekt reduziert, wenn die Urteilenden gezwungen werden, Argumente aufzuführen, die gegen die Höhe des Ankers sprechen.⁴³⁷ Ein Richter muss sich mit den Argumenten der Gegenseite auseinandersetzen, ehe er ein Urteil fällt (BGE 101 Ia 552 f.). Dadurch verringert sich der Einfluss des ersten Antrags auf das Urteil, verschwindet aber nicht vollständig. Die vorgehend erwähnte Studie von MUSSWEILER et al. zeigt, dass der Effekt auch bei Auflistung von dem Ankerwert widersprechenden Argumente nicht vollständig verschwindet und signifikant bleibt.

2. Rechtliche Rahmenbedingungen, die den Einfluss des Ankereffekts mässigen

- 224 Die *rechtlichen* Faktoren, die dem Ankereffekt entgegenwirken, sind mannigfaltig:

⁴³⁴ MUSSWEILER/STRACK/PFEIFFER, FN 403; NORTHCROFT/NEALE, FN 404.

⁴³⁵ Vgl. MUSSWEILER, FN 407.

⁴³⁶ Vgl. RITOV, FN 405.

⁴³⁷ MUSSWEILER/STRACK/PFEIFFER, FN 403, 1146.

- 225 Als erstes ist sicher die europäische Tradition der Kostenverteilung nach Obsiegen und Unterliegen anzuführen. Das Risiko des Überklagens verbietet es, absurd hohe Anträge zu stellen, um das Gericht zu beeinflussen. Immerhin gibt es Rechtsgebiete, in denen die Kostenverlegung nicht mässigend wirkt. Werden Genugtuungsansprüche adhäsionsweise im Strafverfahren geltend gemacht, so werden die Kosten nach dem grundsätzlichen Obsiegen und Unterliegen verteilt, und nicht nach dem betragsmässigen Ergebnis.⁴³⁸ Auch im grundsätzlich kostenlosen Sozialversicherungsprozess besteht kein Risiko des Überklagens. Als wichtigstes Rechtsgebiet aber ist das Strafrecht zu nennen: die Staatsanwaltschaft trägt auch dann keine Kosten, wenn ihr Antrag nicht oder nur zu einem geringen Teil gutgeheissen wurde. Nach dem Gesagten kann man ihr nur raten, einen möglichst hohen Strafantrag zu stellen.
- 226 Zweitens hat die Kenntnis des richtigen, oder zumindest annähernd richtigen, Zielwerts einen mässigenden Einfluss auf den Ankereffekt.⁴³⁹ Der Einfluss des Ankereffekts ist daher geringer, wo es eine gefestigte Praxis gibt – so z.B. bei der Strafzumessung beim Fahren in angetrunkenem Zustand, wo eigentliche (allerdings kantonal unterschiedliche) „Tarife“ bestehen. Ist das Gericht hingegen mit einer Straftat konfrontiert, die es nur selten zu beurteilen hat und zu der es wenig Präjudizien gibt, dann ist die Unsicherheit über die angemessene Strafe grösser und der Einfluss des Ankereffekts stärker. Durch geschicktes Platzieren eines Ankers können Staatsanwälte daher in *cases of first impression* einen Präzedenzfall schaffen, der die zukünftige Strafzumessung für dieses Delikt zu bestimmen vermag.
- 227 Genereller formuliert wirken alle Faktoren gegen den Einfluss des Ankereffekts, die das gerichtliche Ermessen einschränken, weil sie das Wissen über den korrekten Zielwert vergrössern. Dazu gehört die Bindung an – oder zumindest die starke Beachtung von⁴⁴⁰ – Präjudizien. Sie helfen, Wissen über den Zielwert zu generieren, das nicht vom Anker abhängt. Nach Konsultation der Präjudizien durch die Richter wäre der Einfluss des Ankers in der oben dargestellten Studie sicher geringer ausgefallen (allerdings behaupte ich, dass er nicht gänzlich verschwunden wäre). Im Fall der Genugtuung lässt die psychologische Forschung zum Ankereffekt es als angezeigt erscheinen, denjenigen Modellen zur Berechnung der Genugtuung den Vorzug zu geben, die das Gericht an einen objektiven Massstab binden.⁴⁴¹ Die bundesgerichtliche Forderung nach Berücksichtigung aller konkreten Umstände (BGE 112 II 131, 133) dient der Einzelfallgerechtigkeit, erschwert aber die Vergleichbarkeit der Fälle, vergrössert die Unsicherheit über die im konkreten Fall angemessene Genugtuungssumme und führt dazu, dass dem Ankereffekt grosser Einfluss zukommt.

⁴³⁸ Zürcher Praxis, ZR 86 (1987) Nr. 56.

⁴³⁹ MUSSWEILER/STRACK, FN 408, 515.

⁴⁴⁰ Ich will hier nicht der *stare decisis* das Wort reden.

⁴⁴¹ HÜTTE/DUKSCH, FN 430, 21 ff.

III. Ankereffekt in Verhandlungen

- 228 Der Ankereffekt legt es nahe, dass in Verhandlungen diejenige Partei, die das erste Angebot macht und dadurch den Anker setzen kann, einen erheblichen Vorteil genießt. Zahlreiche empirische Studien bestätigen diese Vermutung.⁴⁴²
- 229 KOROBKIN und GUTHRIE zeigen, dass es wahrscheinlicher ist, dass eine Partei ein Vergleichsangebot über \$ 12'000 akzeptiert, wenn das erste Angebot der Gegenseite \$ 2'000 war, als wenn das erste Angebot bereits bei \$ 10'000 lag.⁴⁴³ Möglicherweise stellen sich diejenigen, die ein erstes Angebot über \$ 2'000 bekommen haben, darauf ein, relativ wenig zu erhalten. Ein letztes Angebot über \$ 12'000 erscheint dann vergleichsweise grosszügig. Umgekehrt erscheinen \$ 12'000 als relativ wenig, wenn bereits das erste Angebot bei \$ 10'000 lag. Das erste Angebot verankert m. a. W. die Erwartungen der Teilnehmenden und beeinflusst deren Bereitschaft, ein letztes Angebot zu akzeptieren.⁴⁴⁴
- 230 In einer Studie von ADAM D. GALINSKY und THOMAS MUSSWEILER verhandelten MBA-Studenten über den Verkauf einer fiktiven pharmazeutischen Fabrik. Die Verhandlungspartner erhielten realistische Informationen über den wirtschaftlichen Hintergrund des Deals. Ob der Käufer oder der Verkäufer das erste Gebot abgeben konnte, wurde von den Studienleitern manipuliert.
- 231 Die Verhandlungsergebnisse zeigten einen überwältigenden Vorteil für den Erstbieter. Machte der Käufer das erste Gebot, so einigten sich die Parteien im Schnitt auf einen Preis von \$ 19,7 Mio. Machte hingegen der Verkäufer das erste Gebot, so wurden durchschnittlich \$ 24,8 Mio. als Preis vereinbart.⁴⁴⁵
- 232 Der Ankereffekt wirkt nicht nur auf die Gegenseite, sondern auch auf den Verhandelnden selbst. Wer sich auf sein Wunschergebnis konzentriert, erzielt objektiv bessere Resultate als derjenige, der sich seine beste Alternative zu einer Einigung⁴⁴⁶ vor Augen hält.⁴⁴⁷

⁴⁴² JEROME M. CHERTKOFF/MELINDA CONLEY, Opening Offer and Frequency of Concessions as Bargaining Strategies, *Journal of Personality and Social Psychology* 1967, 181-185; R.M. LIEBERT/W.P. SMITH/J.H. HILL/M. KEIFFER, The effects of information and magnitude of initial offer on interpersonal negotiation, *Journal of Experimental Social Psychology* 1967, 431-441; HENRIK KRISTENSEN/TOMMY GÄRLING, Anchoring Induced Biases in Consumer Price Negotiations, *Journal of Consumer Policy* 2000, 445-460; RITOV, FN 405, 21.

⁴⁴³ KOROBKIN/GUTHRIE, FN 218, 12.

⁴⁴⁴ KOROBKIN/GUTHRIE, FN 218, 12.

⁴⁴⁵ ADAM D. GALINSKY/THOMAS MUSSWEILER, First Offers as Anchors: The Role of Perspective-Taking and Negotiator Focus, *Journal of Personality and Social Psychology* 2001, 657-669.

⁴⁴⁶ Best Alternative To a Negotiated Agreement, BATNA, also derjenige Preis, unter oder über dem die Partei ohne Einigung besser dasteht. Siehe ROGER FISHER/WILLIAM URY/BRUCE PATTON, *Getting to Yes*, 2. Aufl., New York 1991.

⁴⁴⁷ SALLY B. WHITE/MARGARET A. NEALE, The Role of Negotiator Aspirations and Settlement Expectancies in Bargaining Outcomes, *Organizational Behavior and Human Decision Processes* 1994, 303-317, 311; ADAM D. GALINSKY/THOMAS MUSSWEILER/VICTORIA H. MEDVEC, Disconnecting Outcomes and Evaluations: The Role of Negotiator Focus, *Journal of Personality and Social Psychology* 2002, 1131-1140, 1136.

Gleichzeitig ist er aber mit dem erzielten Resultat unzufriedener als derjenige, der sich von Anfang an das Minimalziel vorstellte und vergleichsweise weniger erreichte.⁴⁴⁸

1. Tipps für den gewinnbringenden Einsatz des Ankereffekts in Verhandlungssituationen

233 Die Stärke des Ankereffekts legt es nahe, ihn in Verhandlungssituationen gezielt einzusetzen. Die folgenden Tipps ergeben sich zwanglos aus der dargestellten psychologischen Forschung:

a. Nach den Sternen greifen

234 Verhandelnde sollten sich hohe Ziele setzen.⁴⁴⁹ Sie werden zwar subjektiv vom Ausgang der Verhandlung enttäuschter sein als diejenigen, die sich mit dem Minimum zufriedengeben. Aber sie erzielen objektiv bessere Ergebnisse.

235 Das Risiko hoher Erwartungen soll aber nicht verschwiegen bleiben: neben der Enttäuschung, dass die Ziele nicht erreicht wurden, besteht es vor allem darin, dass allzu hohe Erwartungen zum Scheitern der Verhandlungen führen können, weil sich die Parteien nicht mehr finden.⁴⁵⁰ Ein ausschliesslich konfrontativer Verhandlungsstil ist häufig ineffizient.⁴⁵¹

b. Das erste Angebot machen

236 Verhandelnde sollten versuchen, das erste Gebot abzugeben. Sie verschaffen sich einen grossen Vorteil, weil sie den ersten Anker setzen können, der die von der Gegenseite zum Verhandlungsobjekt abgerufenen Informationen zu ihren Gunsten verzerrt.

c. Nicht zu bescheiden sein

237 Verhandelnde sollten viel verlangen, ohne zu übertreiben (plausible Anker wirken stärker). Das macht sie zwar auch nicht sympathischer, beeinflusst aber das Verhandlungsergebnis

⁴⁴⁸ GALINSKY/MUSSWEILER/MEDVEC, FN 447, 1137.

⁴⁴⁹ DIRK VAN POUCKE/MARC BUELENS, Predicting the Outcome of a Two-Party Price Negotiation: Contribution of Reservation Price, Aspiration Price and Opening Offer, *Journal of Economic Psychology* 2002, 67-76, 68: "higher aspiration prices [...] lead to better outcomes"; WILLIAM URY, *Getting Past No*, New York 1993, 25: "low aspirations tend to be self-fulfilling. What you don't ask for, the other side is unlikely to give you".

⁴⁵⁰ RUSSELL KOROBKIN, Aspirations and Settlement, *Cornell Law Review* 2002, 1-61, 57.

⁴⁵¹ ANDREA KUPFER SCHNEIDER, Shattering Negotiation Myths: Empirical Evidence on the Effectiveness of Negotiation Style, *Harvard Negotiation Law Review* 2002, 143-233.

zu ihren Gunsten. Zudem bleiben ihnen die unguuten Gefühle erspart, die entstehen, wenn das erste Angebot einfach akzeptiert wird...⁴⁵²

2. Was man tun kann, um den Einfluss des Ankereffekts zu verringern

- 238 Eine Erfolg versprechende Strategie, um sich selber gegen die Beeinflussung durch den Anker der Gegenseite zu immunisieren, ist sich auf die Umstände zu konzentrieren, die *gegen* die Plausibilität des Ankers sprechen. Die Suche nach Informationen, die gegen die Angemessenheit des Ankers sprechen, scheint zumindest geeignet, den Ankereffekt in Verhandlungssituationen zu reduzieren.⁴⁵³
- 239 Was soll man tun, wenn die Gefahr besteht, dass ein Dritter sich durch den Anker ungebührlich beeinflussen lässt? Den Richter oder Schiedsrichter auf den Ankereffekt aufmerksam zu machen und ihn zu bitten, sich vom Rechtsbegehren der Gegenseite nicht ungebührlich beeinflussen zu lassen, wird wenig nützen, denn der Ankereffekt wirkt selbst dann, wenn die Urteilenden gewarnt und gebeten wurden, den Anker nicht zu beachten.⁴⁵⁴ Wenn die Gefahr besteht, dass ein Dritter vom Anker beeinflusst wird, muss dieser dazu gebracht werden, sich die gegen den Anker sprechenden Argumente konkret vorzustellen.⁴⁵⁵ Die normale anwaltliche Tätigkeit sollte geeignet sein, beim Richter oder beim Schiedsrichter solche Überlegungen auszulösen.

⁴⁵² ADAM D. GALINSKY/VANESSA L. SEIDEN/PETER H. KIM/VICTORIA H. MEDVEC, The Dissatisfaction of Having Your First Offer Accepted: The Role of Counterfactual Thinking in Negotiations, *Personality and Social Psychology Bulletin* 2002, 271-283.

⁴⁵³ GALINSKY/MUSSWEILER, FN 445, 661.

⁴⁵⁴ WILSON/HOUSTON/ETLING/BREKKE, FN 401, 397.

⁴⁵⁵ MUSSWEILER/STRACK/PFEIFFER, FN 403, 1146.

§ 9 Darstellungseffekt (*framing*)

I. Der Darstellungseffekt im Recht

- 240 Gerichts- und Schiedsverfahren bilden einen natürlichen Rahmen, um finanzielle Aussichten als Gewinne oder Verluste darzustellen. In den meisten Fällen zivilrechtlicher Klagen hat die Klägerin die Wahl, entweder eine gewisse Summe als Vergleich zu akzeptieren oder den Prozess fortzuführen in der Hoffnung, durch ein Urteil mehr zu erhalten. Die Beklagte hingegen wählt zwischen der Alternative, mit Sicherheit einen gewissen Betrag zur sofortigen Erledigung des Streites zu bezahlen und der Chance, zu einer geringeren oder gar keiner Zahlung verurteilt zu werden. Kläger wählen daher zwischen Optionen, die – immer verglichen mit dem Status quo – Gewinne darstellen, während Beklagte zwischen Optionen wählen, die Verluste repräsentieren.⁴⁵⁶ Die Prospect Theory sagt daher voraus, dass Kläger eher den Vergleich – die risikolose Wahl – vorziehen, während Beklagte die Fortführung des Verfahrens, also die riskante Option, wählen (siehe vorne, S. 75 ff.).

A. Laborexperimente

- 241 JEFFREY RACHLINSKI hat diese Voraussage der Prospect Theory in einem einfachen Laborexperiment mit Studierenden überprüft.⁴⁵⁷ Die Hälfte seiner Versuchspersonen – die Kläger – hatte die Wahl, entweder einen Vergleich über \$ 200'000 zu akzeptieren oder den Prozess weiterzuführen und mit einer Wahrscheinlichkeit von 50% \$ 400'000 zu gewinnen. Die andere Hälfte – die Beklagten – hatten die Wahl, für die vergleichsweise Erledigung des Falles \$ 200'000 zu bezahlen oder ein Risiko von 50 % einzugehen, bei Weiterführung des Prozess zur Bezahlung von \$ 400'000 verurteilt zu werden. 77 % der Kläger, aber nur 31 % der Beklagten, wählten den Vergleich, obwohl in beiden Fällen der erwartete Wert der Alternativen identisch ist.⁴⁵⁸
- 242 Weitere Laborexperimente mit komplexeren Sachverhalten bestätigen, dass Beklagte regelmässig risikogeneigt sind, also den Prozess wählen, während Kläger den Vergleich vorziehen.⁴⁵⁹ Ein Experiment von LINDA BABCOCK und Kolleginnen zeigt zudem, dass auch Prozessanwälte mit durchschnittlich zwölf Jahren Berufserfahrung dem Darstellungseffekt unterliegen.⁴⁶⁰

⁴⁵⁶ RACHLINSKI, FN 219, 118 f.

⁴⁵⁷ RACHLINSKI, FN 219, 128.

⁴⁵⁸ RACHLINSKI, a.a.O.

⁴⁵⁹ RUSSEL KOROBKIN/CHRIS GUTHRIE, Psychological Barriers to Litigation Settlement: An Experimental Approach, Michigan Law Review 1994, 107-192, 133 ff.; RACHLINSKI, FN 219, 136 f.

⁴⁶⁰ LINDA BABCOCK/HENRY S. FARBER/CYNTHIA FOBIAN/ELDAR SHAFIR, Forming Beliefs About Adjudicated Outcomes: Perceptions of Risk and Reservation Values, International Review of Law and Economics 1995, 289-303, 296 f.

B. Feldstudien

- 243 Verhalten sich Beklagte auch ausserhalb des Labors – zu ihrem finanziellen Schaden – risikogeneigt? Die Daten von zwei amerikanischen Feldstudien legen diesen Schluss nahe.
- 244 Gemäss dem ökonomischen Modell des Prozessverhaltens ist davon auszugehen, dass Kläger und Beklagte je in der Hälfte der Fälle, die bis zum Urteil durchprozessiert werden, obsiegen.⁴⁶¹ Der Kläger akzeptiert vergleichsweise eine Summe, die der entspricht, die er sich bei positivem Ausgang des Prozesses erhofft, abzüglich der Kosten der Prozessführung.⁴⁶² Der Beklagte andererseits offeriert vergleichsweise eine Summe, die der entspricht, zu der er im schlechtesten Fall verurteilt würde, zuzüglich der Kosten der Prozessführung. Das Vergleichsangebot hängt in diesem Modell also entscheidend von der Einschätzung der Prozesschancen durch die Parteien ab. Da die Kosten des Prozesses zur Bezahlung des Beklagten addiert und von der dem Kläger zugesprochenen Summe subtrahiert werden, gibt es in den allermeisten Fällen eine Zone, innerhalb der beide Parteien besser fahren, wenn sie sich vergleichen. Die Fälle, die bis zum Urteil durchprozessiert werden, sind diejenigen Fälle, in denen eine Partei das zu erwartende Urteil falsch eingeschätzt hat – mit anderen Worten werden in der Regel nur die schwierigen Streitigkeiten, deren Ausgang kaum prognostizierbar ist, bis zum bitteren Ende ausgefochten.⁴⁶³ Unter den Annahmen, dass beide Parteien risikoneutral sind, über die gleichen Informationen verfügen und gleich kompetent vertreten sind, müsste sich in 50 % der Fälle der Kläger verschätzen, in den anderen 50 % der Beklagte. Daher impliziert die Selektions-Hypothese – dass nur die schwierigen Fälle durchprozessiert werden – eine 50 %ige Chance des Obsiegens für jede Partei.⁴⁶⁴
- 245 Diese Vorhersage des ökonomischen Modells von PRIEST/KLEIN lässt sich empirisch überprüfen. SAMUEL R. GROSS und KENT D. SYVERUD unternahmen diesen Versuch, basierend auf Daten von 529 Fällen, die von kalifornischen Geschworenengerichten entschieden wurden.⁴⁶⁵ Auf den ersten Blick schienen die aggregierten Daten die Hypothese des ökonomischen Modells zu bestätigen: in 51,4 % der Fälle gewann der Kläger, in 48,6 % der Beklagte.⁴⁶⁶ Aber nicht jedes Mal, wenn dem Kläger eine geringe Summe zugesprochen wird, kann er ernsthaft als erfolgreich bezeichnet werden: die durch das Urteil zugesprochene Summe muss zumindest über dem letzten Vergleichsangebot des Beklagten liegen, ansonsten war der (erhebliche) Aufwand für den Prozess vergeblich.

⁴⁶¹ GEORGE L. PRIEST/BENJAMIN KLEIN, The Selection of Disputes for Litigation, *Journal of Legal Studies* 1984, 1-55.

⁴⁶² Gemäss amerikanischen Zivilprozessrecht werden die Prozesskosten in der Regel auch bei Obsiegen nicht erstattet. Da die Parteientschädigung gemäss Anwaltstarif in der Schweiz zumindest bei geringen bis mittleren Streitwerten die tatsächlichen Anwaltskosten nicht deckt, ist das Modell von PRIEST/KLEIN durchaus geeignet, auch schweizerische Prozesse zu modellieren.

⁴⁶³ PRIEST/KLEIN, FN 461, 13.

⁴⁶⁴ PRIEST/KLEIN, FN 461, 5.

⁴⁶⁵ SAMUEL R. GROSS/KENT D. SYVERUD, Getting to No: A Study of Settlement Negotiations and the Selection of Cases for Trial, *Michigan Law Review* 1991, 319-393. Die Daten stammen aus *Jury Verdict Weekly*; *Jury Verdict Weekly* erfasst rund 90% aller Geschworenengerichtsurteile in Kalifornien, siehe GROSS/SYVERUD, 332.

⁴⁶⁶ GROSS/SYVERUD, FN 465, 333.

GROSS/SYVERUD nahmen als Massstab daher die durch das Urteil zugesprochene Summe abzüglich des letzten Angebots des Beklagten. Würden sich die Beklagten risikoneutral verhalten, müsste die durch Urteil zugesprochene Summe mal über, mal unter dem letzten Vergleichsangebot der Beklagten liegen; die Masszahl folglich im Durchschnitt bei null liegen.⁴⁶⁷

- 246 Die den Klägern zugesprochene Summe lag aber in allen Fällen, in denen entsprechende Daten erhoben werden konnten, über Null;⁴⁶⁸ das von den Beklagten gemachte Vergleichsangebot war mit anderen Worten ungenügend. Dieses Resultat lässt sich durch den Darstellungseffekt erklären.⁴⁶⁹ Wenn Beklagte risikogeneigt sind, dann sind sie an einer vergleichweisen Beilegung des Streites weniger interessiert und werden Vergleichsangebote machen, die unter dem erwarteten Wert des Urteils liegen.⁴⁷⁰
- 247 Für die Situation in der Schweiz besonders interessant ist, dass in Handelssachen, wo auch in den USA beide Parteien in der Regel ihre Anwälte nach Aufwand entschädigen und selber bezahlen müssen, der Vorteil der Kläger besonders gross war.⁴⁷¹ Die Kläger in Handelssachen bekamen in 80% der Fälle mehr zugesprochen als das letzte Vergleichsangebot des Beklagten.⁴⁷² Die Kläger verlangten oft weniger, als ihnen schlussendlich zugesprochen wurde, und die Beklagten boten häufig gar nichts an, um den Prozess durch Vergleich zu erledigen. GROSS/SYVERUD schreiben dies der Tatsache zu, dass der Kläger in Handelssachen sehr viel eigenes Geld investieren muss, ehe er zu einem Urteil kommt. Beklagte, die dies wissen, spekulieren damit, dass der Kläger dieses Geld nicht aufbringen wird. Wenn der Kläger aber bereit und fähig ist, das Geld aufzuwerfen, dann hat er in der Regel einen sehr guten *case*, und gewinnt entsprechend in 80 % der Fälle.⁴⁷³
- 248 In einer weiteren Feldstudie analysiert RACHLINSKI 656 ebenfalls von kalifornischen Gerichten entschiedene Fälle.⁴⁷⁴ Auch er vergleicht die letzten Vergleichsangebote von Kläger und Beklagten mit den letztlich zugesprochenen Summen. Ihre Risikoneigung kommt die Beklagten teuer zu stehen, wenn der Streit durch Urteil entschieden wird. In den von RACHLINSKI analysierten Zivilstreitigkeiten wären die Beklagten im Schnitt \$ 66'106 besser gefahren, wenn sie sich verglichen hätten, während die Kläger vom Urteil im gleichen Umfang profitierten.⁴⁷⁵ Dies gilt auch dann, wenn beide Parteien juristische Personen sind (die sich theoretisch rationaler verhalten sollten). Die unterlassene Einigung

⁴⁶⁷ Genau genommen sogar unter Null; denn letztes Angebot des Beklagten = erwarteter Prozessausgang + Prozesskosten. Da die Prozesskosten nicht zugesprochen werden, müsste der Prozessausgang im Schnitt tiefer sein als das letzte Angebot des Beklagten.

⁴⁶⁸ GROSS/SYVERUD, FN 465, 354 (für Personenschäden), 357 (für Verkehrsunfälle), 369 (für Handelssachen), 374 f. (für Arbeitsstreitigkeiten).

⁴⁶⁹ GROSS/SYVERUD, FN 465, 371, schreiben es der (fehlgeschlagenen) Verhandlungsstrategie der Beklagten zu.

⁴⁷⁰ RACHLINSKI, FN 219, 150.

⁴⁷¹ In *personal injury cases* wird hingegen meist ein Erfolgshonorar vereinbart, was zu anderen Anreizen für strategisches Verhandeln führt.

⁴⁷² GROSS/SYVERUD, FN 465, 367.

⁴⁷³ GROSS/SYVERUD, FN 465, 370.

⁴⁷⁴ RACHLINSKI, FN 219, 152 ff.

⁴⁷⁵ RACHLINSKI, FN 219, 155.

kostete die beklagten juristischen Personen im Schnitt \$ 550'282 pro Streitfall – ohne Anwaltskosten.⁴⁷⁶

- 249 Auch in der Schweiz gewinnt der Kläger weit häufiger als der Beklagte. Im Kanton Bern obsiegte der Kläger in arbeitsrechtlichen Streitigkeiten im Jahr 2003 in 48 von 69 durch Urteil entschiedenen Fällen vollständig (69 %), in weiteren 14 Fällen wurde teilweise zugunsten des Klägers entschieden, und nur in sieben Fällen vollständig zugunsten des Beklagten.⁴⁷⁷ Die aargauischen Bezirksgerichte haben 2002 von 375 durch Urteil entschiedenen Zivilsachen 238 vollständig gutgeheissen (63 %), 74 teilweise gutgeheissen und nur 63 vollständig abgewiesen.⁴⁷⁸ Im Kanton Graubünden haben die Zivilkammern der Bezirksgerichte im Geschäftsjahr 2003 von 170 durch Urteil erledigten Klagen 80 gutgeheissen, weitere 36 teilweise gutgeheissen und 52 abgewiesen.⁴⁷⁹ Auch Klagen vor Handelsgericht – wo Kläger wie Beklagte in der Regel juristische Personen sind – werden häufiger gutgeheissen als abgewiesen. Am Handelsgericht Aargau wurden 2002 von 64 durch Urteil entschiedenen Klagen 27 vollständig und weitere 23 teilweise gutgeheissen (78 %), aber nur 14 abgewiesen.⁴⁸⁰ Die 50:50 Hypothese des klassischen ökonomischen Modells lässt sich auch in der Schweiz nicht belegen.
- 250 Die tatsächlichen Daten stimmen jedoch gut mit den Voraussagen der Prospect Theory überein. Die Kläger, weil sie risikoscheu sind, sind zu schnell bereit, den Streit vergleichsweise zu erledigen, während Beklagte wegen ihrer Risikobereitschaft selbst dann auf einem Urteil bestehen, wenn die voraussichtlich durch das Urteil zugesprochene Summe über dem Vergleichsangebot der Gegenseite liegt. Kläger sollten daher ermutigt werden, weniger nachgiebig zu sein, während Beklagten verstärkt zu einem Vergleich zu raten ist.⁴⁸¹

II. Unterliegen auch Richter dem Darstellungseffekt?

- 251 Die meisten Zivilverfahren werden in der Schweiz nicht durch Urteil, sondern durch Vergleich erledigt. Am Arbeitsgericht Zürich wurden 2003 von 1'205 Verfahren 916 durch Vergleich, Anerkennung oder Rückzug erledigt (76 %); durch Urteil wurden nur 219 (18

⁴⁷⁶ RACHLINSKI, FN 219, 157. Diese Verluste im Urteilsfall werden wahrscheinlich kompensiert durch Gewinne der Beklagten in den verglichenen Fällen, weil die Kläger wegen ihrer Risikoscheu zu tiefe Vergleichsangebote akzeptieren, siehe GROSS/SYVERUD, FN 465, 371. Wegen der fehlenden empirischen Daten über verglichene Fälle ist es nicht möglich, diese Annahme zu überprüfen.

⁴⁷⁷ Verwaltungsbericht des Obergerichts des Kantons Bern, Statistiken 2003 der Untersuchungsrichterämter und der Gerichtskreise, Tafel IV (erhältlich unter www.be.ch/OG/ -> Publikationen; besucht am 29. Januar 2005).

⁴⁷⁸ Obergericht des Kantons Aargau, Rechenschaftsbericht Justizbehörden 2001/2002, Anhang 3, S. 7 (erhältlich unter www.ag.ch/obergericht/de/pub/angebote/rechenschaftsberichte.htm; besucht am 29. Januar 2005). Diese Zahlen sind möglicherweise durch den hohen Anteil an Scheidungssachen (die in der Regel gutgeheissen werden), verzerrt.

⁴⁷⁹ Jahresbericht 2003 des Kantonsgerichts Graubündens, Tabelle 7b 1 (erhältlich unter www.kg-gr.ch/ -> Aktuelles/Infos; besucht am 29. Januar 2005). Die Zahlen für das Vorjahr sind sehr ähnlich: 154 durch Urteil erledigte Verfahren, davon 71 vollständige Guttheissung, 44 Teilguttheissung, 37 Abweisung.

⁴⁸⁰ Obergericht des Kantons Aargau, FN 478, Anhang 1, S. 11.

⁴⁸¹ GUTHRIE/RACHLINSKI/WISTRICH, FN 20, 798.

%) entschieden.⁴⁸² Für das Handelsgericht Zürich, wo praxisgemäss nach dem ersten Schriftenwechsel eine Vergleichsverhandlung unter Vorsitz des Instruktionsrichters durchgeführt wird, sind die Zahlen ähnlich.⁴⁸³ URS EGLI kommt in seiner Dissertation zum Schluss, dass im Kanton Zürich rund 70-80 % aller vergleichsfähigen – d.h. der Dispositionsmaxime unterstehenden – Verfahren durch Vergleich erledigt werden.⁴⁸⁴

- 252 In vielen Kantonen sind Vergleichsverhandlungen unter Leitung des Instruktionsrichters üblich. Der Richter kann dabei, wie praktische Erfahrung und wissenschaftliche Studien⁴⁸⁵ belegen, die Parteien erheblich beeinflussen. Wenn Richter dem gleichen Darstellungseffekt wie die Parteien unterliegen, dann fordern sie möglicherweise die Kläger auf, weniger zu akzeptieren, als angemessen wäre, und üben andererseits nicht genügend Druck auf den Beklagten aus, ein angemessenes Vergleichsangebot zu machen.⁴⁸⁶

A. Amerikanische Studie von Guthrie und Kollegen

- 253 CHRIS GUTHRIE und Kollegen haben deshalb untersucht, ob amerikanische *Magistrate Court Judges* ebenfalls dem Darstellungseffekt unterliegen. Sie legten 163 Richtern einen fiktiven Fall zur Beurteilung vor, bei dem sich zwei Firmen darum stritten, ob die eine die Urheberrechte der anderen verletzt habe. Rechtlich sei der Fall einfach, aber es stellten sich schwierige Beweisfragen, und zurzeit liesse sich nur sagen, dass eine Chance von 50 % bestehe, dass der Kläger gewinne und die eingeklagten \$ 200'000 erhalte. Verliere er, bekomme er nichts. Die Anwaltskosten, um den Prozess bis zum Ende durchzuziehen, betrügen \$ 50'000 pro Partei, und es bestehe keine Möglichkeit, diese Kosten von der anderen Partei ersetzt zu bekommen.⁴⁸⁷
- 254 Die eine Hälfte der Richter beurteilte den Sachverhalt aus der Perspektive des Klägers und wurde gefragt, ob der Kläger ein Angebot des Beklagten, den Fall für \$ 60'000 zu vergleichen, akzeptieren solle. Die andere Hälfte der Richter wurde gefragt, ob sie dem Beklagten empfehlen würden, dem Kläger \$ 140'000 zu bezahlen, um den Fall vergleichsweise zu erledigen.⁴⁸⁸
- 255 Der Kläger steht somit vor einer Wahl, \$ 60'000 auf sicher zu akzeptieren oder ein „Spiel“ mit einem erwarteten Wert von \$ 50'000 ($0,5 \cdot \$ 200'000 \text{ Urteil} + 0,5 \cdot \$ 0 \text{ Urteil} - \$ 50'000 \text{ Anwaltskosten}$) zu spielen. Der Beklagte andererseits steht vor der Wahl, einen sicheren Verlust von \$ 140'000 zu realisieren oder ein Spiel mit einem erwarteten „Wert“ von - \$ 150'000 zu spielen ($0,5 \cdot - \$ 200'000 \text{ Urteil} + 0,5 \cdot \$ 0 \text{ Urteil} - \$ 50'000 \text{ Anwaltskosten}$). In beiden Fällen ist der Vergleich daher die bessere, und einzig rationale, Wahl.

⁴⁸² Rechenschaftsbericht des Obergerichts des Kantons Zürich 2003, 100 (70 Verfahren wurden „anderweitig“ erledigt).

⁴⁸³ A. a. O., 154.

⁴⁸⁴ URS EGLI, Vergleichsdruck im Zivilprozess, Berlin 1996, 63.

⁴⁸⁵ EGLI, a. a. O., 104 ff.

⁴⁸⁶ GUTHRIE/RACHLINSKI/WISTRICH, FN 20, 798.

⁴⁸⁷ GUTHRIE/RACHLINSKI/WISTRICH, FN 20, 796.

⁴⁸⁸ GUTHRIE/RACHLINSKI/WISTRICH, FN 20, 796.

- 256 40 % der Richter, die den Fall aus der Perspektive des Klägers beurteilten, empfahlen die Annahme des Vergleichs gegenüber nur 25 % der Richter, die den Fall aus Sicht des Beklagten beurteilten. Der Unterschied zwischen den Gruppen ist statistisch signifikant.⁴⁸⁹ Richter scheinen, wie Parteien, vom Darstellungseffekt beeinflusst zu werden.

III. Eigene Studie zum Darstellungseffekt

A. Frage zum Darstellungseffekt

- 257 Den Richterinnen und Richtern wurde folgender Sachverhalt vorgelegt (2003):

Sie sind ReferentIn in einem Prozess zwischen zwei Baufirmen. Die Klägerin hat die Beklagte auf Zahlung eines Gewinnanteils von Fr. 150'000 aus einem gemeinsamen Bauprojekt verklagt. Beide Baufirmen sind mittelgrosse Unternehmen mit einem Jahresumsatz von rund CHF 10 Mio. pro Jahr. Sie werden von fähigen Anwälten vertreten, die Ihnen nicht persönlich bekannt sind.

Wird der Fall bis zu einem Urteil durchprozessiert, so hat jede Partei mit eigenen Anwaltskosten von Fr. 30'000 zu rechnen. Die Prozessentschädigung gemäss Gebührenverordnung beträgt rund Fr. 15'000; selbst bei einem vollständigen Obsiegen erhält die obsiegende Partei daher Fr. 15'000 ihrer Anwaltskosten nicht ersetzt. Die von der unterliegenden Partei zu bezahlenden Gerichtskosten betragen Fr. 10'000.

Der Fall ist rechtlich einfach. Die Höhe des geschuldeten Betrages ist nicht umstritten, nur der Anspruch der Klägerin wird bestritten. Der Ausgang des Prozesses hängt einzig von der Aussage eines Schlüsselzeugen ab. Sie glauben, dass eine 50 %ige Chance besteht, dass die Klägerin die vollen Fr. 150'000 erhält und eine 50 %ige Chance, dass die Klägerin Fr. 0 erhält.

- 258 Die eine Hälfte der befragten Richterinnen beurteilten den Fall aus Sicht der Klägerin („Klägerin/Gewinn-Gruppe“). Ihnen wurde die folgende Frage gestellt:

An der Referentenaudienz bietet die Beklagte der Klägerin eine Zahlung von Fr. 50'000 für den Rückzug der Klage an. Die Gerichtskosten würden von beiden Parteien je zur Hälfte übernommen und jede Partei trüge ihre eigenen Anwaltskosten. Für ihren Anteil an den Gerichtskosten und die eigenen Anwaltskosten muss jede Partei mit Fr. 10'000 rechnen.

Glauben Sie, dass die Klägerin die Zahlung von Fr. 50'000 akzeptieren sollte, um den Prozess zu erledigen?

- 259 Die Klägerin hatte also die Wahl zwischen sicheren Fr. 40'000⁴⁹⁰ oder einem „Spiel“, bei dem sie eine 50 % Chance hat, Fr. 135'000⁴⁹¹ zu gewinnen und eine 50 % Chance, Fr. 55'000⁴⁹² zu verlieren. Der erwartete Wert des Spiels beträgt daher für die Klägerin Fr. 40'000 ($0,5 \cdot \text{Fr. } 135'000 - 0,5 \cdot \text{Fr. } 55'000$).

- 260 Der anderen Hälfte der Richter wurde erklärt, dass die Klägerin der Beklagten anbiete, die Klage gegen eine Zahlung von Fr. 100'000 zurückzuziehen.⁴⁹³ Diese Richter wurden gefragt, ob die Beklagte Fr. 100'000 zahlen sollte, um den Prozess zu erledigen („Beklagte/Verlust-Gruppe“).

⁴⁸⁹ GUTHRIE/RACHLINSKI/WISTRICH, FN 20, 797.

⁴⁹⁰ Fr. 50'000 Vergleichszahlung – Fr. 10'000 Anwalts- und Gerichtskosten = Fr. 40'000.

⁴⁹¹ Fr. 150'000 Urteil + Fr. 15'000 Parteientschädigung – Fr. 30'000 Anwaltskosten = Fr. 135'000.

⁴⁹² – Fr. 30'000 Anwaltskosten – Fr. 15'000 Parteientschädigung – Fr. 10'000 Gerichtskosten = – Fr. 55'000.

⁴⁹³ Wiederum unter Wettschlagung der Anwalts- und hälftiger Teilung der Gerichtskosten.

Besonderer Teil

- 261 Die Beklagte hat also die Wahl, entweder Fr. 110'000⁴⁹⁴ sofort zu bezahlen oder mit einem Risiko von 50 % entweder Fr. 205'000⁴⁹⁵ oder Fr. 15'000⁴⁹⁶ zu bezahlen. Der negative Wert des Spiels beträgt für die Beklagte daher $- \text{Fr.} 110'000 (0.5 \cdot - \text{Fr.} 205'000 + 0.5 \cdot - \text{Fr.} 15'000)$.
- 262 Es ist wichtig zu sehen, dass die Alternativen wirtschaftlich identisch sind. Die Parteien streiten sich über die Aufteilung eines fixen „Kuchens“, nämlich des Gewinns aus dem gemeinsamen Projekt. Beide Parteien stehen vor der Wahl, entweder Fr. 40'000 des gemeinsamen Gewinns zu behalten oder ein Spiel mit dem erwarteten Wert von Fr. 40'000 zu spielen. Die Beklagte ist im Besitz des gemeinsamen Gewinns von Fr. 150'000. Nach Bezahlung von Fr. 100'000 an die Klägerin und der eigenen Anwaltskosten von Fr. 10'000 bleiben ihr Fr. 40'000. Umgekehrt erhält die Klägerin, die nichts vom gemeinsamen Gewinn besitzt, bei Annahme des Vergleichsangebots Fr. 50'000. Nachdem sie ihre Anwaltskosten von Fr. 10'000 bezahlt hat, bleiben ihr folglich ebenfalls Fr. 40'000 vom gemeinsamen Gewinn. Einer risikoneutralen Person ist es gleichgültig, ob sie die sichere oder die riskante Alternative wählt. Eine risikoscheue Person – wegen des abnehmenden Grenznutzens des Vermögens der Normalfall – zieht die sichere Alternative vor. Auf jeden Fall sollte es keinen Unterschied geben, der davon abhängt, wer die Fr. 150'000 besitzt.

B. Resultate

- 263 In der Gruppe der Richter, die den Fall aus der Sicht der Klägerin beurteilten, empfahlen 56,6 % der Richter (43 von 76), die angebotene Zahlung von Fr. 50'000 zu akzeptieren und den Streit zu vergleichen. Nur 43 % der Richter (43 von 100), die den Fall aus der Perspektive der Beklagten beurteilten, empfahlen jedoch der Beklagten, Fr. 100'000 zu bezahlen, um den Fall abzuschliessen. Die Wahrscheinlichkeit, dass es keinen Unterschied zwischen den beiden Gruppen gibt und die unterschiedliche Verteilung der Antworten auf Zufall zurückzuführen ist, beträgt 7,4 %.⁴⁹⁷ Der Unterschied ist daher knapp nicht signifikant auf dem 5 %-Niveau.

	Aus Sicht des Klägers	Aus Sicht des Beklagten
Schweiz	56,6 %	43 %
USA	39,8 %	25 %

Tabelle 8: Vergleichsempfehlung

- 264 Die schweizerischen Richter sind vergleichsfreundlicher als ihre amerikanischen Kollegen. Während 39,8 % der amerikanischen Richter, die den Fall aus der Perspektive der Kläger beurteilten, zum Vergleich rieten, empfahlen 56,6 % der schweizerischen Richter, den

⁴⁹⁴ $-\text{Fr.} 100'000 - \text{Fr.} 10'000 \text{ Anwalts- und Gerichtskosten} = -\text{Fr.} 110'000$.

⁴⁹⁵ $-\text{Fr.} 150'000 \text{ Urteil} - \text{Fr.} 30'000 \text{ Anwaltskosten} - \text{Fr.} 15'000 \text{ Parteientschädigung} - \text{Fr.} 10'000 \text{ Gerichtskosten} = -\text{Fr.} 205'000$.

⁴⁹⁶ $\text{Fr.} 0 \text{ Urteil} - \text{Fr.} 30'000 \text{ Anwaltskosten} + \text{Fr.} 15'000 \text{ Parteientschädigung} = -\text{Fr.} 15'000$.

⁴⁹⁷ $\chi^2 = 3,187$; $p = 0,074$; Likelihood-Quotient = 3,195.

Streit zu vergleichen.⁴⁹⁸ Dabei lag der erwartete Wert des Urteils bei der amerikanischen Studie von GUTHRIE/RACHLINSKI/WISTRICH bei \$ 50'000, während der Vergleich einen sicheren Gewinn von \$ 60'000 versprach; m .a. W. hätte ein rationaler risikoneutraler Kläger den Vergleich wählen *müssen* (während es ihm in meiner Studie egal sein konnte). Bei den Richtern, die den Fall aus der Perspektive der Beklagten beurteilten, lagen die Werte bei 25 % für einen Vergleich bei den amerikanischen Richtern verglichen mit 43 % der schweizerischen Richter.⁴⁹⁹

C. Diskussion

- 265 Obwohl das Vergleichsangebot, das den Klägern und den Beklagten gemacht wurde, wirtschaftlich identisch ist, raten Richter, die den Fall aus Sicht der Beklagten beurteilen, seltener zu einem Vergleich als diejenigen, die den Fall aus klägerischer Sicht betrachten.
- 266 Wenn Richter dem gleichen Darstellungseffekt unterliegen wie die Parteien, werden sie der Klägerin eher zum Vergleich raten als der Beklagten.⁵⁰⁰ Es sollte aber gerade umgekehrt sein: es sind die Beklagten, die zu ihrem eigenen Nutzen vermehrt dazu gebracht werden sollten, ein vernünftiges Vergleichsangebot zu machen. Sie gehen zu hohe Risiken ein und verlieren in der Folge die meisten Prozesse.
- 267 Ein Richter, der den Darstellungseffekt kennt, kann seine Position dazu nutzen, die Sichtweise der Parteien zu beeinflussen. Die meisten Entscheidungen lassen sich sowohl als Wahl zwischen Verlustrisiken als auch als Wahl zwischen Gewinnchancen darstellen.⁵⁰¹ Im Beispielfall des Prozesses zwischen den beiden Bauunternehmen kann ein Richter versuchen, der Beklagten klar machen, dass sie bei einem Vergleich nicht nur Fr. 110'000 bezahlen muss, sondern auch Fr. 40'000 des gemeinsamen Gewinns behalten kann. Dadurch lässt sich die Entscheidung als Wahl zwischen Gewinnen statt als Wahl zwischen unterschiedlich hohen Verlusten darstellen. Entsprechend sollte die Beklagte eher geneigt sein, die risikolose Alternative, also den Vergleich, zu wählen.

IV. Prospect Theory und Klagen mit geringen Erfolgsaussichten

- 268 Der empfundene Wert einer riskanten Wahl ergibt sich gemäss der Prospect Theory sowohl aus der Wertfunktion als auch aus der Gewichtung der Wahrscheinlichkeit. Geringe Wahrscheinlichkeiten werden subjektiv überschätzt (siehe vorne S. 60 f.).
- 269 Für Klagen mit geringen Aussichten auf Erfolg bedeutet dies, dass die Überschätzung der geringen Erfolgschance durch den Kläger dazu führen kann, dass er sich – anders als für „gewöhnliche“ Klagen, deren Erfolgsaussichten in der Nähe von 50 % liegen – risikogeneigt verhält.

⁴⁹⁸ GUTHRIE/RACHLINSKI/WISTRICH, FN 20, 797. Der Unterschied zwischen den beiden Kläger/Gewinn-Gruppen ist signifikant; $\chi^2 = 4,498$, $p = 0,034$.

⁴⁹⁹ Der Unterschied ist hier allerdings nicht signifikant, $\chi^2 = 1,191$, $p = 0,275$.

⁵⁰⁰ GUTHRIE/RACHLINSKI/WISTRICH, FN 20, 796.

⁵⁰¹ TVERSKY/KAHNEMAN, FN 325, 268 f.

- 270 Auch diese Vorhersage der Prospect Theory wurde empirisch überprüft. GUTHRIE fragte seine Versuchspersonen, ob sie lieber einen Vergleich abschliessen und \$ 50 erhalten oder den Prozess bis zum Urteil durchziehen möchten, wobei eine Chance von einem Prozent bestünde, dass sie den Prozess gewinnen und \$ 5'000 erhielten. Den Versuchspersonen in der Rolle der Beklagten wurde angeboten, den Fall für \$ 50 zu vergleichen oder das Risiko von einem Prozent einzugehen, bei einem Urteil \$ 5'000 zu verlieren.⁵⁰² 62 % der Versuchspersonen in der Rolle des Klägers wählten den Prozess, also die riskante Wahl, während 84 % der Versuchspersonen in der Rolle der Beklagten den Vergleich vorzogen.⁵⁰³ Die Umkehr der Risiko-Präferenz bei geringer Wahrscheinlichkeit trat also, wie von der Prospect Theory vorhergesagt, auf.⁵⁰⁴
- 271 „*Frivolous Litigation*“, d.h. Klagen mit geringen Erfolgsaussichten, werden in den USA als grosses Problem empfunden.⁵⁰⁵ Sie belasten das ohnehin bereits überlastete Justizsystem zusätzlich und führen zu hohen Anwaltskosten bei den Beklagten, die diese wegen der amerikanischen Regel (jede Partei trägt ihre eigenen Kosten) selbst bei vollständigem Obsiegen nicht ersetzt bekommen. Die Analyse der Gründe, die zu leichtfertigen Klagen führen, ist daher von einigem Interesse.
- 272 In der Schweiz halte ich den Erkenntnisgewinn aus dieser Analyse für begrenzt, weil frivole Klagen kaum ein ernsthaftes Problem sind. Erstens setzen sie voraus, dass Kläger – wenn auch mit einer geringen Wahrscheinlichkeit – im Erfolgsfall sehr hohe Summen zugesprochen bekommen. In der Schweiz wird kein Strafschadenersatz zugesprochen, sondern nur der tatsächlich erlittene Schaden (inkl. entgangenem Gewinn, wobei die Beweisforderungen sehr hoch sind), ersetzt. Nicht finanzielle Schäden werden nur äusserst zurückhaltend monetär abgegolten (siehe vorne S. 83). Es fehlt daher am Anreiz, frivole Klagen einzureichen.
- 273 Zweitens muss die unterliegende Partei in der Schweiz die Anwaltskosten der Gegenpartei ersetzen. Wenn auch die zugesprochenen Summen in der Praxis bei geringen bis mittlerem Streitwert nicht ausreichen, die tatsächlichen Anwaltskosten zu bezahlen, so ist die Aussicht für den Kläger, im Falle seines Unterliegens der Gegenseite eine Prozessentschädigung leisten zu müssen, in der Regel abschreckend genug, um ihn von der leichtfertigen Einreichung einer Klage abzuhalten. Auch GUTHRIE geht davon aus, dass frivole Klagen unter der „britischen Regel“ der Kostenverlegung (entspricht der schweizerischen Regel) kaum vorkommen.⁵⁰⁶
- 274 Klagen mit äusserst geringen Erfolgsaussichten werden in der Schweiz daher erfahrungsgemäss nur von eigentlichen Querulanten eingereicht, die sich nicht von Überlegungen zum erwarteten Wert der Klage leiten lassen.

⁵⁰² CHRIS GUTHRIE, *Framing Frivolous Litigation: A Psychological Theory*, University of Chicago Law Review 2000, 163-216, 188.

⁵⁰³ GUTHRIE, a. a. O.

⁵⁰⁴ GUTHRIE, a. a. O., 190, führte auch noch ein Experiment mit einem komplexeren Sachverhalt durch, das die Ergebnisse des ersten Experimentes bestätigte.

⁵⁰⁵ Nachweise bei GUTHRIE, a. a. O., 163, Fn. 1.

⁵⁰⁶ GUTHRIE, a. a. O., 213.

V. Prospect Theory und Steuerhinterziehung

- 275 Steuerhinterziehung ist ein riskantes Spiel: wird der Hinterziehende erwischt, bezahlt er nebst Nachsteuern und Zinsen eine Busse, die das Dreifache der hinterzogenen Steuer betragen kann (Art. 175 Abs. 2 Bundesgesetz über die direkte Bundessteuer, DBG). Steuerhinterziehung verspricht aber natürlich auch Ersparnisse, wenn sie nicht entdeckt wird.
- 276 Die Prospect Theory sagt voraus, dass das Risikoverhalten des Steuerpflichtigen unterschiedlich ist, je nachdem, ob er die zu entrichtende Steuer als Gewinn oder Verlust empfindet. Wie aber kann man eine Steuer als Gewinn empfinden?
- 277 In Ländern, in denen direkte Steuern an der Quelle erhoben werden – z.B. in den USA und Deutschland die Einkommenssteuer von Lohnempfängern – kann der Steuerpflichtige in die Lage kommen, dass er nach Einreichung der Steuererklärung eine Rückerstattung der zuviel bezogenen Steuer erhält. Diese Rückerstattung wird er subjektiv als Gewinn empfinden, weil er über das Geld zurzeit, d.h. im Status quo, nicht verfügen kann.⁵⁰⁷ Wer andererseits – wie der schweizerische Lohnempfänger in der Schweiz⁵⁰⁸ – die Steuern aus dem Geld bezahlen muss, das er bereits auf seinem Konto hat und über das er frei verfügen kann, empfindet die Steuer als Verlust.
- 278 Für den Steuerpflichtigen, der die Steuer aus dem Geld bezahlen muss, über das er verfügen kann, erscheint Steuerhinterziehung daher als riskante Wahl zwischen zwei Verlusten: entweder einem sicherem Verlust bei vollständiger Deklaration aller steuerpflichtigen Einkünfte oder einem geringeren Verlust bei Hinterziehung eines Teils der Einkünfte, wobei das Risiko besteht, dass der Verlust bei Entdeckung der Hinterziehung umso grösser ausfällt. Gemäss der Prospect Theory wird er sich risikogeneigt verhalten, um den sicheren Verlust zu vermeiden, und deshalb eher Steuern hinterziehen.⁵⁰⁹
- 279 Andererseits wählt der Steuerpflichtige, der eine Rückerstattung der bereits bezogenen Steuer erwarten kann, zwischen zwei unterschiedlich hohen Gewinnen: einem sicheren kleinen Gewinn, wenn er vollständig deklariert, und einem grösseren Gewinn, sprich einer grösseren Rückerstattung, wenn er einen Teil der Einkünfte nicht deklariert. Er wird sich gemäss der Prospect Theory risikoscheu verhalten.
- 280 Sowohl empirische wie auch experimentelle Studien bestätigen diese Voraussagen der Prospect Theory.⁵¹⁰ In Ländern, in denen die Steuer an der Quelle abgeführt wird, kommt es weniger häufig zu Falschdeklarationen. In den USA deklarierten 96 % der Steuerpflichtigen, die eine Rückerstattung von \$ 1'000 oder mehr zu erwarten hatten, ihre Einkünfte korrekt, gegenüber nur 89 % derjenigen, die mit einer Nachzahlung von mehr als \$ 1'000

⁵⁰⁷ KENNETH W. SMITH/KARYL A. KINSEY, Understanding Taxpaying Behavior: A Conceptual Framework with Implications for Research, Law and Society Review 1987, 639-663.

⁵⁰⁸ Bei Ausländern ohne fremdenpolizeiliche Niederlassungsbewilligung wird die Steuer an der Quelle erhoben, d.h. direkt vom Arbeitgeber abgeführt, Art. 83 Abs. 1, Art. 88 DBG.

⁵⁰⁹ RUSSEL H. WEIGEL/DICK J. HESSING/HENK ELFFERS, Tax Evasion Research: A Critical Appraisal and Theoretical Model, Journal of Economic Psychology 1987, 215-235.

⁵¹⁰ Nachweise bei ERICH KIRCHLER/BORIS MACIEJOVSKY, Steuermoral und Steuerhinterziehung, in: DIETER FREY/LUTZ VON ROSENSTIEL (Hrsg.), Enzyklopädie der Psychologie – Wirtschaftspsychologie, Göttingen 2002.

rechnen müssen.⁵¹¹ Bei den Geschäftseinkünften ist der Unterschied noch ausgeprägter: 95 % vollständige Deklarationen bei Steuerpflichtigen, die eine Rückerstattung erwarten gegenüber nur 70 % bei Steuerpflichtigen, die eine Nachzahlung erwarten.⁵¹²

- 281 ELFFERS und HESSING schlagen deshalb vor, vermehrt Vorauszahlungen zu verlangen, um die Steuermoral zu verbessern.⁵¹³ ELIZABETH LOFTUS hat dem amerikanischen Steueramt bereits 1985 empfohlen, die Lohnabzüge zu erhöhen, so dass sich mehr Personen vor Einreichung der Steuererklärung in der Situation befinden, zwischen „Gewinnen“ zu wählen.⁵¹⁴
- 282 Ein Steuersystem, bei dem ein Grossteil des Steueraufkommens nicht an der Quelle, sondern beim Steuerpflichtigen erhoben wird, hat auch Vorteile. Direkte Steuern, die erhoben werden, nachdem der Steuerpflichtige bereits Verfügungsgewalt über die steuerbaren Einkünfte erhalten haben, schmerzen besonders, weil Verluste mehr wehtun als Gewinne erfreuen (*loss aversion*). Die Steuerpflichtigen werden sich daher gegen Steuererhöhungen besonders wehren. MCCAFFREY argumentiert deshalb, und belegt dies mit entsprechenden Zahlen, dass „versteckte“ Steuern, d.h. Steuern, die erhoben werden, bevor der Steuerpflichtige die Verfügungsgewalt über das steuerbare Einkommen erlangt, überproportional angestiegen sind. Nach MCCAFFREY ist dies deshalb der Fall, weil Steuererhöhungen bei den versteckten Steuern weniger Widerstand erwächst, wobei er dem Staat nicht unterstellt, die Steuerlast absichtlich auf solche Steuern zu verlagern, sondern es bei der Hypothese bewenden lässt, dass dies gleichsam „natürlich“ oder „evolutionär“ geschieht, weil die Steuerlast dem Weg des geringsten Widerstandes folgt.⁵¹⁵
- 283 Der Schweizer Staat, der (noch) einen relativ grossen Anteil der gesamten Steuerlast durch direkte Einkommenssteuern erhebt, die vom Steuerpflichtigen aus Einkünften bezahlt werden, über die er frei verfügen kann,⁵¹⁶ wird gemäss dieser Sicht mehr Schwierigkeiten haben, höhere Steuern durchzusetzen, als ein Staat, der sein Steueraufkommen aus indirekten Steuern oder aus Quellensteuern bezieht. Ähnliche Argumente, allerdings ohne Bezug zur Prospect Theory von KAHNEMAN/TVERSKY, wurden bereits von WALTER

⁵¹¹ HENRY S. J. ROBBEN/PAUL WEBLEY/RUSSELL H. WEIGEL/KARL-ERIK WARNERYD/KARYL A. KINSEY/DICK J. HESSING/FRANCISCO ALVIRA MARTIN/HENK ELFFERS/ RICHARD WAHLUND/LUK VAN LANGENHOVE, Decision frame and opportunity as determinants of tax cheating: An international experimental study, *Journal of Economic Psychology* 1990, 341-364, 346.

⁵¹² A. a. O.

⁵¹³ HENK ELFFERS/DICK J. HESSING, Influencing the Prospects of Tax Evasion, *Journal of Economic Psychology* 1997, 289-304.

⁵¹⁴ ELIZABETH F. LOFTUS, To File, Perchance To Cheat, *Psychology Today*, April 1985, 34 ff., 38, zitiert nach CHRIS GUTHRIE, Prospect Theory, Risk Preference, And The Law, *Northwestern University Law Review* 2003, 1115-1163, 1145.

⁵¹⁵ MCCAFFREY, FN 194, 404.

⁵¹⁶ Der Anteil der direkten Steuern am gesamten Steueraufkommen in der Schweiz betrug 2001 68,9 %, verglichen mit 56 % im Durchschnitt der OECD Staaten, siehe Eidgenössische Steuerverwaltung, Abteilung Steuerstatistik und Dokumentation, Direkte vs. Indirekte Steuern: Einige Bemerkungen zum Steuermix in der Schweiz, Bern 2003, 1 (erhältlich unter www.estv.admin.ch/data/d/paket/tax_mix.pdf, besucht am 29. Januar 2005).

HETTICH und STANLEY WINER vorgebracht.⁵¹⁷ Es überrascht daher nicht, dass die Eidgenössische Steuerverwaltung – mit zahlreichen guten Gründen natürlich – für eine Erhöhung des Anteils der indirekten Steuern am Gesamtsteueraufkommen eintritt.⁵¹⁸

VI. Status quo Bias und Vertragsverhandlungen

- 284 Die traditionelle ökonomische Analyse von Vertragsverhandlungen beruht auf der fundamentalen Prämisse, dass die Präferenz einer Partei für eine bestimmte Vertragsklausel endogen ist und nicht von der Darstellung in der Verhandlung abhängt. Externe Faktoren können durchaus beeinflussen, welche Vertragsklausel letztlich vereinbart wird – beispielsweise, weil eine an und für sich weniger bevorzugte Vertragsklausel bereits in einem Standardvertrag enthalten ist, und es aufwändig und teuer wäre, den Vertrag neu zu schreiben – aber sie beeinflussen nicht, welche Klausel in Abwesenheit der externen Faktoren bevorzugt würde. Externe Faktoren können, mit anderen Worten, das *Ergebnis* einer Verhandlung beeinflussen, nicht aber die *Präferenzen* der Parteien für ein bestimmtes Ergebnis.⁵¹⁹
- 285 In einem funktionierenden Markt sollte das Ergebnis der Verhandlungen daher in der Regel die Präferenzen der Parteien spiegeln. Da in den USA in Gliedstaaten, in denen die Kündigung eines Arbeitsvertrages ohne Begründung zulässig ist („*termination at will*“), nur 15 % der Arbeitsverträge abweichend von der dispositiven Regel eine Klausel enthalten, dass eine Kündigung nur aus zureichendem Grund („*just cause*“) zulässig sei, schliessen Vertreter der traditionellen ökonomischen Analyse des Rechts, dass die grosse Mehrheit der Vertragsschliessenden Kündigungen ohne Grund präferieren.⁵²⁰
- 286 Der *Status quo Bias* hingegen besagt, dass die Präferenz einer Partei für eine Klausel davon abhängt, ob die Klausel dem Status quo entspricht. Als Status quo Bias wird die Neigung bezeichnet, eine Wahl zu treffen, die den Ist-Zustand erhält oder leicht verbessert.⁵²¹ Die Bevorzugung des Status quo wird mit zwei Faktoren erklärt: der Abneigung gegenüber Verlusten (*loss aversion*), die bereits vorne, S. 58 f. erläutert wurde, und der Bevorzugung von Unterlassungen (Unterlassungseffekt; *omission bias*).⁵²²

⁵¹⁷ WALTER HETTICH/STANLEY L. WINER, A Positive Model of Tax Structure, *Journal of Public Economics* 1984, 67-87.

⁵¹⁸ Eidgenössische Steuerverwaltung, FN 516, 3.

⁵¹⁹ RUSSELL KOROBKIN, Behavioral Economics, Contract Formation, and Contract Law, in: SUNSTEIN (Hrsg.), 116-143, 118.

⁵²⁰ J. HOULT VERKERKE, An Empirical Perspective on Indefinite Term Employment Contracts: Resolving the Just Cause Debate, *Wisconsin Law Review* 1995, 837 ff., 867, zitiert nach RUSSELL KOROBKIN, The Status Quo Bias and Contract Default Rules, *Cornell Law Review* 1998, 608-687, 669 f.

⁵²¹ TVERSKY/KAHNEMAN, FN 501, 257.

⁵²² MAURICE SCHWEITZER: Disentangling Status Quo and Omission Effects: An Experimental Analysis, *Organizational Behavior and Human Decision Making Processes* 2002, 457-476.

A. Demonstration des Status quo Biases bei Präferenzen für Vertragsklauseln

- 287 SUNSTEIN hat einer Gruppe von (amerikanischen) Jura-Studierenden folgendes Szenario vorgelegt: Angenommen, sie haben eine Stelle bei einer Anwaltskanzlei angenommen. Das Jahressalär beträgt \$ 120'000 und der gesetzliche Anspruch auf Urlaub zwei Wochen pro Jahr. Der Hälfte der Versuchspersonen wurde gesagt, die Kanzlei gewähre freiwillig zwei Wochen Urlaub zusätzlich im Jahr. Diese Versuchspersonen wurden gefragt, wie viel mehr ihnen die Kanzlei bezahlen müsse, damit sie auf ihren Anspruch auf die zusätzlichen Urlaubstage verzichten. Der anderen Hälfte der Versuchspersonen wurde mitgeteilt, dass die Kanzlei nur den gesetzlichen Anspruch gewähre, aber mit sich reden liesse, gegen eine entsprechende Lohnreduktion zwei Wochen Urlaub zusätzlich zu gewähren, wobei keine weiteren Nachteile irgendwelcher Art zu befürchten seien. Die Versuchspersonen in dieser Gruppe wurden gefragt, welche Lohneinbusse sie in Kauf nehmen würden, um zwei Wochen Urlaub zusätzlich zu erhalten. Sie boten im Median \$ 6'000 an, während die Personen in der anderen Gruppe im Median \$ 13'000 verlangten, um auf die zwei Wochen Urlaub zu verzichten. Vier Studierende, deren Antworten bei der Berechnung des Medians nicht berücksichtigt wurden, meinten, sie würden für kein Geld auf der Welt auf ihre zwei Wochen zusätzlichen Urlaub verzichten.⁵²³

B. Die Bevorzugung des Status quo als Folge der Abneigung gegen Handlung

- 288 Die Bevorzugung des Status quo kann als Folge der Abneigung gegen Handlungen, die meist notwendig sind, um den Status quo zu verändern, gesehen werden. WILLIAM SAMUELSON und SAMUEL ZECKHAUSER baten ihre Versuchspersonen, zwischen vier verschiedenen Anlagemöglichkeiten für ihr gerade eben geerbtes Vermögen zu wählen – eine riskante Aktienanlage, einer weniger riskanten Anlage in Aktien, einer Schatzanweisung und Gemeindefinanzen. Je einem Fünftel der Versuchspersonen wurde gesagt, dass das geerbte Vermögen bereits in einer der vier Anlagen investiert sei (dem letzten Fünftel, der Kontrollgruppe, dass das Vermögen in liquider Form vorliege).⁵²⁴ Es zeigte sich, dass jede der vier Gruppen eine andere Anlageform bevorzugte – nämlich diejenige, in die das Geld bereits investiert war.⁵²⁵
- 289 Ein Experiment von KAHNEMAN kann einen ersten Anhaltspunkt dafür geben, warum Menschen Nichtstun – und damit meist den Status quo – bevorzugen. KAHNEMAN erklärte seinen Versuchspersonen, dass Paul, der Aktien der Gesellschaft A besitzt, sich überlegt habe, sie zu verkaufen und stattdessen Aktien der Gesellschaft B zu kaufen, sich aber letztlich dagegen entschieden habe. Später erfährt Paul, dass er \$ 1'200 mehr verdient hätte, wenn er den Wechsel vollzogen hätte. George andererseits besass Aktien der Gesellschaft B und entschied sich, sie zu verkaufen und in Aktien der Gesellschaft A zu investieren. Später erfährt er, dass er \$ 1'200 mehr verdient hätte, wenn er die Aktien der Gesellschaft B behalten hätte. Wer wird seinen Entscheid mehr bedauern? 92 % der Versuchspersonen

⁵²³ CASS R. SUNSTEIN, Switching the Default Rule, New York University Law Review 2002, 106-134, 113 f.

⁵²⁴ WILLIAM SAMUELSON/RICHARD ZECKHAUSER, Status Quo Bias in Decision Making, Journal of Risk and Uncertainty 1988, 7-59, 12.

⁵²⁵ SAMUELSON/ZECKHAUSER, FN 524, 17, Tabelle 1c.

waren der Meinung, dass George – der gehandelt hat – seine Entscheidung mehr bedauern würde.⁵²⁶

- 290 Dass Menschen grösseres Bedauern antizipieren, wenn eine Handlung zu einem Misserfolg führt als wenn eine Unterlassung zum gleichen Misserfolg führt, ist ein in zahlreichen Experimenten belegtes Phänomen.⁵²⁷ Das mag daran liegen, dass man eine Handlung eher als Ursache für einen Erfolg wahrnimmt und sich daher stärker verantwortlich für Erfolg oder Misserfolg einer Handlung als einer Unterlassung fühlt.⁵²⁸
- 291 Der Unterschied in antizipiertem Bedauern kann die Bevorzugung des Status quo erklären: da die Beibehaltung des Status quo meist mit Nichtstun/Unterlassung einhergeht, wird die Angst vor dem Bedauern, eine falsche Wahl getroffen zu haben, dazu führen, dass man „sicherheitshalber“ alles beim Alten lässt.
- 292 ILANA RITOV und JONATHAN BARON haben versucht, Einfluss von Unterlassung und Beibehaltung des Status quos, die in den früheren Experimenten meist nicht zu trennen sind, auf das antizipierte Bedauern zu unterscheiden. Sie haben deshalb das Paul/George Beispiel von KAHNEMAN abgewandelt. Henry, dem Aktien der Gesellschaft A gehören, wird von seinem Vermögensverwalter vorgeschlagen, stattdessen in Aktien der Gesellschaft B zu investieren. Henry lehnt ab und behält somit den Status quo durch aktives Tun bei. Frank hingegen, der Aktien der Gesellschaft B besitzt, wird von seinem Vermögensverwalter gefragt, ob er etwas dagegen habe, stattdessen in Aktien der Gesellschaft A zu investieren. Frank widerspricht nicht und verändert damit seine Vermögensanlage ohne aktive Handlung. Frank und Henry erfahren wiederum, dass sie bei gegenteiliger Entscheidung \$ 1'200 mehr verdient hätten. Die Versuchspersonen wurden gefragt, wer von den beiden sich wohl schlechter fühle.
- 293 Wenn die Änderung des Status quo die Hauptursache des Bedauerns ist, müsste Frank grösseres Bedauern empfinden, da sich bei ihm der Status quo geändert hat. Ist hingegen die Unterscheidung zwischen Handlung und Unterlassung massgebend, dann sollte Henry grösseres Bedauern empfinden, denn er hat aktiv gehandelt, um den Status quo beizubehalten. Die Versuchspersonen bei RITOV/BARON gaben an, dass sich Henry schlechter fühlen würde als Frank.⁵²⁹ RITOV/BARON schliessen daraus – und aus zahlreichen weiteren Experimenten – dass die Bevorzugung des Status quo in erster Linie auf die unterschiedlich stark bedauerten Folgen von Handlungen und Unterlassungen zurückzuführen ist.⁵³⁰ SCHWEITZER bestätigt den Einfluss des Unterlassungseffekts auf die Bevorzugung des

⁵²⁶ DANIEL KAHNEMAN, Varieties of Counterfactual Thinking, in: NEAL J. ROESE/JAMES M. OLSON, What Might Have Been: The Social Role of Counterfactual Thinking, New York 1995, 375-396, 388.

⁵²⁷ THOMAS GILOVICH/VICTORIA H. MEDVEC, The Experience of Regret: What, When, and Why, Psychological Review 1995, 379-395, 380: “perhaps the clearest and most frequently replicated finding in the entire literature on counterfactual thinking”.

⁵²⁸ JOHANNA H. KORDES-DE VAAL, Intention and the omission bias: Omissions perceived as nondecisions, Acta Psychologica 1996, 161-172, 169: “the basis of the omission bias is a difference in perceived causality – probably caused by the fact that commissions are more salient than omissions – making the outcome of an omission appear less intended than the outcome of a commission”; ähnlich auch schon ILANA RITOV/JONATHAN BARON, Status-quo and Omission Bias, Journal of Risk and Uncertainty 1992, 49-61, 60.

⁵²⁹ RITOV/BARON, FN 528, 53.

⁵³⁰ RITOV/BARON, FN 528, 59

Status quo, widerspricht aber RITOV/BARON insofern, als der Wunsch nach Beibehaltung des Status quos an und für sich gemäss seinen Experimenten ebenfalls einen Einfluss auf die Stärke des Effekts hat.⁵³¹ Für den Juristen entscheidend ist letztlich, dass die Bevorzugung des Status quos, so sie mit Unterlassung einhergeht, ein robustes und in zahlreichen Experimenten bestätigtes Phänomen ist.

C. Dispositives Recht, Allgemeine Vertragsbedingungen und Status quo

- 294 Die Annahme, dass Parteien das dispositive Recht als Status quo und Abweichungen davon als Veränderung des Status quo ansehen, liegt nahe. Zwingend ist sie nicht. Dispositives Vertragsrecht verschafft niemandem eine Stellung, die vergleichbar ist mit derjenigen, die ein absolutes Recht verschafft. Den Vorteil einer dispositiven vertragsrechtlichen Norm genießt nur, wer eine andere Partei dazu bringt, einen Vertrag abzuschliessen, auf den die Norm anwendbar ist und der die Norm nicht abändert. Die Vorteile, die sich aus einer dispositiven vertragsrechtlichen Norm ergeben können, sind immer Folge der Kooperation zweier Parteien.⁵³²

1. Empirische Studien von Russell Korobkin

- 295 RUSSELL KOROBKIN hat empirisch überprüft, ob Parteien dispositives Vertragsrecht als „Besitz“ empfinden, was dazu führt, dass sie die dispositive Regel, weil Status quo, bevorzugen.
- 296 Zwei Unternehmen, „NextDay“ (eine Frachtführerin), und die Geschenk AG, eine Absenderin von Frachtgut, verhandeln über den Abschluss eines Rahmenvertrages. Der Hälfte der Versuchspersonen wurde gesagt, dass das dispositive Recht die Haftung des Frachtführers bei Verzug auf „vernünftigerweise voraussehbare“ Schäden beschränkt (beschränkte Haftung), der anderen Hälfte, dass gemäss dispositivem Recht der Frachtführer für den gesamten Verzugsschaden, voraussehbar oder nicht, haftet (volle Haftung). Die Übernahme der vollen Haftung würde NextDay zwischen \$ 0 - \$ 10 pro Paket kosten.⁵³³
- 297 Die Versuchspersonen in der „beschränkte Haftung“ Gruppe wurden gefragt, was die Geschenk AG NextDay bezahlen muss, damit NextDay die volle Haftung übernimmt; die Versuchspersonen in der „volle Haftung“ Gruppe, welche Preisreduktion NextDay akzeptieren soll, wenn die Geschenke AG dafür auf die volle Haftung verzichtet und sich mit der beschränkten Haftung zufrieden gibt. Die Personen in der „beschränkte Haftung“ Gruppe verlangten im Schnitt \$ 6,96 pro Paket, um die volle Haftung zu übernehmen, während die Personen in der „volle Haftung“ Gruppe nur \$ 4,46 Preisnachlass gewährten, wenn die Haftung eingeschränkt wurde.⁵³⁴

⁵³¹ SCHWEITZER, FN 522, 470.

⁵³² KOROBKIN, FN 520, 632.

⁵³³ KOROBKIN, FN 519, 122.

⁵³⁴ KOROBKIN, FN 519, 122.

- 298 Was geschieht, wenn Allgemeine Geschäftsbedingungen eine andere Regel als das dispositive Vertragsrecht vorsehen? Wird die dispositive Regel oder die AGB bevorzugt? In einem weiteren Experiment wurde den Versuchspersonen zusätzlich gesagt, dass die Parteien vereinbart hätten, auf der Basis des Standardvertrags eines Branchenverbandes zu verhandeln. Die Konditionen „beschränkte Haftung“ und „volle Haftung“ wurden gekreuzt; d.h. wenn die dispositive gesetzliche Regel eine beschränkte Haftung vorsah, sah der Standardvertrag die volle Haftung vor und umgekehrt. Wiederum wurden die Versuchspersonen gefragt, was NextDay für die Übernahme der vollen Haftung verlangen, resp. für die Reduktion der Haftung gewähren soll. Für die Übernahme der vollen Haftung wurde im Schnitt \$ 7,36 pro Paket verlangt, aber nur ein Preisnachlass von \$ 4,10 gewährt.⁵³⁵
- 299 Die Parteien bevorzugen offenbar immer diejenige Regel, die gilt, wenn sie nichts tun, um sie zu ändern. Dies lässt sich dadurch erklären, dass das antizipierte Bedauern grösser ist, wenn eine Handlung zum Misserfolg führt, als wenn eine Unterlassung zum Misserfolg führt.
- 300 Ob antizipiertes Bedauern tatsächlich der Grund für die Bevorzugung des Status quo – sei er dispositives Gesetzesrecht oder AGB-Klausel – ist, lässt sich überprüfen. Vor zukünftigem Bedauern kann man sich nur dann fürchten, wenn man die negativen Folgen einer Vertragsklausel im Zeitpunkt der Verhandlung nicht mit Sicherheit abschätzen kann (was der Normalfall sein dürfte). Kennt man die Vor- und Nachteile einer Vertragsklausel im Zeitpunkt der Verhandlung mit Sicherheit, bleibt kein Raum für zukünftiges Bedauern.
- 301 In einem weiteren Experiment sagte KOROBKIN daher seinen Versuchspersonen, dass die Versicherungsgesellschaft von NextDay das zusätzliche Risiko der vollen Haftung für einen fixen Betrag übernehmen würde, womit über die Folgen der Haftungsübernahme keine Ungewissheit mehr bestand. In diesem Szenario war der Unterschied zwischen den beiden Gruppen nicht mehr signifikant, was die Hypothese, dass die Bevorzugung des Status quo auf antizipiertes Bedauern zurückzuführen ist, bestätigt.⁵³⁶ Die Hypothese wird ebenfalls durch Experimente bestätigt, die zeigen, dass der Unterlassungseffekt stärker ist, wenn man weiss, dass man die Folgen der Handlung oder Unterlassung erfahren wird.⁵³⁷ Wenn man die Folgen nie kennen wird, kann man auch kein Bedauern antizipieren – was ich nicht wissen werde, wird mich nicht heiss machen.
- 302 Natürlich gibt es auch nach traditioneller ökonomischer Analyse des Rechts Gründe, die für die Beibehaltung einer Klausel eines Standardvertrags (bei konstanten Präferenzen) sprechen. Erstens sind die Transaktionskosten höher, wenn ein Vertrag geändert werden muss. Zweitens geben die Verhandelnden unter Umständen private Informationen preis, wenn sie eine von der üblichen Norm abweichende Regelung der Risikoverteilung verlan-

⁵³⁵ KOROBKIN, FN 519, 125.

⁵³⁶ RUSSELL KOROBKIN, Inertia and Preference in Contract Negotiation: The Psychological Power of Default Rules and Form Terms, *Vanderbilt Law Review* 1998, 1583-1651, 1625. Allerdings besteht nach wie vor ein Unterschied, nur ist er nicht mehr signifikant, was die Annahme bestätigt, dass antizipiertes Bedauern nicht der *einzige* Grund für den Status quo Bias ist.

⁵³⁷ ILANA RITOV/JONATHAN BARON, Outcome Knowledge, Regret, and Omission Bias, *Organizational Behavior and Human Decision Processes* 1995, 119-127.

gen. In dem obigen Beispiel könnte das Beharren der Geschenke AG auf der Übernahme der vollen Haftung durch die NextDay signalisieren, dass die Geschenke AG weiss, dass das Risiko unvorhersehbarer Schäden grösser ist als üblich. Schliesslich können Netzwerkeffekte dafür sprechen, eine allgemein übliche Formulierung beizubehalten. Bei Standardklauseln gibt es oft bereits Rechtsprechung zur Auslegung und dementsprechend weniger Unklarheiten bei der Auslegung, was helfen kann, Streit zu vermeiden. Obwohl diese Erklärungen durchaus plausibel sind, können sie die Resultate des Experiments von KOROBKIN nicht erklären, denn die Versuchsbedingungen wurden sorgfältig kontrolliert, um den Einfluss dieser Variablen auszuschliessen.⁵³⁸ Die Experimente von KOROBKIN lassen nur den Schluss zu, dass die Tatsache allein, dass eine Regel ohne aktives Tun zum Vertragsinhalt wird, genügt, die Präferenzen der Parteien für diese Regel zu ändern.

2. Schlussfolgerung für den Praktiker

- 303 Die Experimente zeigen, dass es simplistisch ist, von der Tatsache, dass eine Klausel in einer Mehrzahl von Verträgen vorkommt, darauf zu schliessen, dass diese Klausel den Präferenzen der Parteien entspricht und (daher) effizient ist.⁵³⁹ SUNSTEIN schlägt daher beispielsweise vor, das dispositive Arbeitsrecht, das in den USA praktisch ausschliesslich den Arbeitgeber bevorzugt, zugunsten der Arbeitnehmer auszugestalten.⁵⁴⁰ Der – aus europäischer Sicht bescheidene – Vorschlag ist schlau, weil er die Vertreter der traditionellen ökonomischen Analyse des Rechts in ein Dilemma bringt: wenn sie sich dagegen wehren, geben sie implizit zu, dass die Prämisse der konstanten Präferenzen nicht stimmt oder – wenn die Vertreter der BLE Recht haben und sich die Präferenzen ändern – dass es ihnen weniger um die Effizienz an und für sich als um den Vorteil der Arbeitgeber geht. Akzeptieren sie den Vorschlag, hat SUNSTEIN sein Ziel erreicht, den Arbeitnehmern eine bessere Rechtsposition zu verschaffen.
- 304 Das Beharren auf dem Status quo bedeutet aber auch, dass der Verhandelnde, der den Status quo bestimmen kann, einen klaren Vorteil genießt.⁵⁴¹ Wer auch immer den ersten Entwurf eines Vertrags macht, ist im Vorteil.⁵⁴² Dies deckt sich mit dem Ratschlag erfahrener Anwälte, den ersten Entwurf eines Vertrages wenn immer möglich selbst zu schreiben. Die traditionelle Begründung dafür ist aber eine andere: traditionellerweise wird der Vorteil des ersten Entwurfs damit begründet, (i) dass die Kosten der Gegenpartei höher sind, wenn sie den Vertrag ändern muss (ii) dass die Gegenpartei private Informationen preisgeben muss, wenn sie die Änderung des Entwurfs begründen muss oder (iii) dass sie oder ihre Berater sich nicht die Zeit nehmen, den Entwurf so genau zu analysieren, dass sie alle negativen Konsequenzen erkennen.
- 305 Die psychologische Vorliebe für Unterlassungen hingegen bedeutet, dass sich die *Präferenz* für eine Vertragsklausel ändert, je nachdem ob sie durch Handlung oder Unter-

⁵³⁸ KOROBKIN, FN 519, 126 ff.

⁵³⁹ KOROBKIN, FN 519, 137.

⁵⁴⁰ SUNSTEIN, FN 523, 133.

⁵⁴¹ KOROBKIN, FN 519, 136.

⁵⁴² KOROBKIN, FN 536, 1627.

lassung zum Vertragsbestandteil wird. Nicht externe Faktoren, sondern die den Verhandeln innewohnende Reaktion auf den Status quo führt dazu, dass die Parteien den Status quo bevorzugen.⁵⁴³

⁵⁴³ KOROBKIN, FN 519, 136.

§ 10 Unterlassungseffekt (*omission bias*)

- 306 Dass man lieber nicht handelt, wenn die Folgen der Handlung unsicher sind, scheint damit zusammenzuhängen, dass man sich für die (möglicherweise negativen) Folgen einer Handlung mehr verantwortlich fühlt als für die gleichen Folgen einer Unterlassung.⁵⁴⁴ Gilt dies auch, wenn man die Handlung oder Unterlassung eines anderen beurteilt, d.h. beurteilen wir eine Handlung eines anderen, die zu einem Misserfolg geführt hat, negativer als eine Unterlassung, die zum identischen Misserfolg geführt hat?

I. Psychologische Studien zur moralischen Bewertung von Handlungen und vorsätzlichen Unterlassungen

- 307 MARK SPRANCA, ELISA MINSK und JONATHAN BARON sind der erwähnten Frage nachgegangen.⁵⁴⁵ Sie haben ihren Versuchspersonen unter anderem den folgenden Sachverhalt vorgelegt und sie gebeten, die Unmoral der Handlung oder Unterlassung auf einer Skala von 0 (nicht unmoralisch) bis 100 (äusserst unmoralisch) zu bewerten: In dem Tennisclub, in dem John spielt, findet jährlich ein Turnier statt, das John als bester Spieler des Clubs regelmässig gewinnt. Dieses Jahr jedoch hat die Klubleitung Ivan Lendl als Gaststar eingeladen. John weiss, dass er gegen Ivan in Normalform verlieren wird. Am Abend vor dem Final findet ein Abendessen statt, an dem die Spieler teilnehmen. John erfährt, dass Ivan allergisch auf Cayenne-Pfeffer reagiert und weiss im Gegensatz zu Ivan, dass das House-Dressing des Club-Restaurants Cayenne-Pfeffer enthält. Er denkt sich, dass Ivan wahrscheinlich die ganze Nacht nicht wird schlafen können, wenn er das House-Dressing isst und am nächsten Tag dementsprechend in schlechter Verfassung sein wird, weshalb er (John) eine Chance hat, gegen Ivan zu gewinnen.
- 308 Den Versuchspersonen in der „Handlung“ Gruppe wurde gesagt, dass John, bevor sich Ivan für eine Sauce entscheidet, das House-Dressing empfiehlt. Hätte John nichts gesagt, hätte sich Ivan für die (pfefferfreie) italienische Sauce entschieden. Den Versuchspersonen in der „Unterlassung“ Gruppe wurde gesagt, dass Ivan das House-Dressing bestellt, ehe John etwas gesagt hat. John sitzt neben Ivan und ist sich bewusst, dass Ivan eine andere Sauce bestellt hätte, wenn er ihn vor dem Cayenne-Pfeffer gewarnt hätte. In beiden Fällen erleidet Ivan eine allergische Reaktion und John gewinnt das Spiel.⁵⁴⁶
- 309 Ob Handlung und Unterlassung in diesem Fall wirklich als gleichwertig zu bewerten ist, ist eine umstrittene Frage. Das schweizerische Recht verneint sie bekanntlich (näheres hinten, S. 111 ff.). Manche Philosophen vertreten die Auffassung, dass die Unterscheidung dort, wo der Vorsatz sowohl bei Handlung wie bei Unterlassung gleichermassen vorliegt, nicht

⁵⁴⁴ Siehe die in FN 528 zitierten Autorinnen.

⁵⁴⁵ MARK SPRANCA/ELISA MINSK/JONATHAN BARON, *Omission and Commission in Judgment and Choice*, *Journal of Experimental Social Psychology* 1991, 76-105.

⁵⁴⁶ SPRANCA/MINSK/BARON, FN 545, 80.

haltbar ist,⁵⁴⁷ während andere darauf beharren, dass ein relevanter Unterschied besteht.⁵⁴⁸ Für Vertreter einer konsequentialistischen Ethik, die die Moral einer Handlung oder Unterlassung an den beabsichtigten Konsequenzen misst, kann es ohnehin keinen Unterschied zwischen Handlung und Unterlassung geben, wenn die beabsichtigten Folgen der beiden identisch sind.⁵⁴⁹ Nach der Theorie der rationalen Entscheidung (*rational choice theory*) ist diejenige Alternative vorzuziehen, die den grössten erwarteten Nutzen hat, unabhängig davon, ob es sich um eine Handlung oder Unterlassung handelt.

- 310 Psychologisch gesehen besteht aber offenbar ein Unterschied zwischen Handlung und Unterlassung: 65 % der Versuchspersonen von SPRANCA et al. bewerteten die Handlung als unmoralischer als die vorsätzliche Unterlassung, nur eine einzige die Unterlassung als unmoralischer als die Handlung.⁵⁵⁰
- 311 In den meisten Fällen ist die Unterscheidung zwischen Handlung und Unterlassung sinnvoll: die Folgen von Handlungen sind meist beabsichtigt, die Folgen von Unterlassungen werden häufig nicht angestrebt und sind unbeabsichtigt. Oft lässt sich daher von der Handlung, nicht aber von der Unterlassung, auf den Vorsatz schliessen. SUNSTEIN nennt dies eine „moralische Urteilsheuristik“, die oft richtig ist, aber durch Übergeneralisierung – d.h. Anwendung auf Fälle, auf die sie nicht passt – zu Fehlurteilen führen kann.⁵⁵¹ Es handelt sich um ein Beispiel der von KAHNEMAN/FREDERICK beschriebenen Attribut-Substitution, die von ihnen als Ursache der Urteilsheuristiken gesehen wird: das schwer erkennbare Zielattribut „Vorsatz/kein Vorsatz“ wird durch das leichter erkennbare heuristische Attribut „Handlung/Unterlassung“ ersetzt.⁵⁵²
- 312 Wenn diese These stimmt, müsste die Unterscheidung zwischen Handlungen und Unterlassung verschwinden, wenn der Vorsatz der handelnden und nicht handelnden Personen gleich beurteilt wird. SPRANCA et al. berichten, dass dies nicht zutrifft, d.h. auch wenn der Vorsatz gleich bewertet wird, wird die Handlung als unmoralischer als die Unterlassung bewertet. Ein Drittel der Versuchspersonen, die den Vorsatz von John im Tennis-Szenario bei Handlung und Unterlassung identisch sahen, bewerteten die Handlung

⁵⁴⁷ JONATHAN BENNETT, *Morality and Consequences*, in: STERLING McMURRIN (Hrsg.), *The Tanner Lectures on Human Values*, Bd. 2, Salt Lake City 1981, 45-116; ebenso CASS R. SUNSTEIN, *Moral Heuristics and Moral Framing*, *Minnesota Law Review* 2004, 1556-1597, 1582: “There is no morally relevant difference between the two [sc. Handlungen und Unterlassungen]; the act-omission distinction makes a difference apparent or even clear when it is not real”; für DIETER BIRNBACHER, *Tun und Unterlassen*, Stuttgart 1995, 127 ist die Unterscheidung von Tun und Unterlassen *als solche* moralisch nicht relevant; eine unterschiedliche Bewertung kann aber durch die grössere wahrgenommene Bedrohung durch Handlungen gerechtfertigt werden, a. a. O., 206. BIRNBACHER beruft sich dabei ausdrücklich auf die sozialpsychologische Forschung zum *omission bias*.

⁵⁴⁸ Zur Kontroverse der Aktiv-Passiv-Unterscheidung bei der Sterbehilfe mit zahlreichen Hinweisen MARKUS ZIMMERMANN-ACKLIN, *Euthanasie*, 2. Aufl., Freiburg 2002, 223 ff. ZIMMERMANN-ACKLIN selbst ist der Ansicht, dass sich die Unterscheidung unter Berücksichtigung ausschliesslich intrinsischer Faktoren kaum begründen lässt, unter Berücksichtigung moralpsychologischer Aspekte aber plausibel ist.

⁵⁴⁹ PETER SINGER, *Praktische Ethik*, Stuttgart 1984, 204 ff. (deutsche Übersetzung der englischen Originalausgabe von 1979); siehe auch JOEL FEINBERG, *The Moral Limits of Criminal Law*, Vol. 1: *Harm to Others*, New York 1984, 171.

⁵⁵⁰ SPRANCA/MINSK/BARON, FN 545, 82.

⁵⁵¹ SUNSTEIN, FN 547, 582 f.; siehe auch SPRANCA/MINSK/BARON, FN 545, 79.

⁵⁵² Zur Attributsubstitution als Erklärung für Urteilsheuristiken siehe KAHNEMAN/ FREDERICK, FN 245, 53 f.

trotzdem als unmoralischer.⁵⁵³ SPRANCA et al. führen dies darauf zurück, dass eine Handlung eher als Ursache wahrgenommen wird, was von den Versuchspersonen beispielsweise damit begründet wird, dass die negativen Folgen der Unterlassung auch dann eingetreten wären, wenn John gar nicht anwesend gewesen wäre.⁵⁵⁴

- 313 KORDES-DE VAAL versucht in einer Serie von Experimenten den Einfluss von empfundener Verursachung, Verantwortung und auf die Folgen gerichtetem Vorsatz auf den Unterlassungseffekt zu unterscheiden.⁵⁵⁵ Sie kommt zum Schluss, dass die unterschiedlich empfundene Verursachung, ohne gleichzeitig wahrgenommene unterschiedliche Absicht, nicht genügt, den Unterlassungseffekt zu bewirken.⁵⁵⁶ Dies stützt die Hypothese der moralischen Urteilsheuristik, gemäss der von Handlung auf Vorsatz und von Unterlassung auf fehlenden Vorsatz geschlossen wird. KORDES-DE VAAL stellt ebenfalls fest, dass die Bewertungen von „Ursache“ und „Verantwortung“ hoch korrelieren, was darauf hindeutet, dass sich die Bedeutung der beiden Konzepte in den Augen der Versuchspersonen kaum unterscheidet.⁵⁵⁷
- 314 Das schweizerische Recht sieht, vereinfacht gesagt, eine strafrechtliche Verantwortung für den durch vorsätzliche Unterlassung herbeigeführten Erfolg in der Regel nur dann vor, wenn zwischen Täter und Geschädigtem eine qualifiziert enge Beziehung besteht. Entspricht diese Unterscheidung zwischen einer vorsätzlichen Unterlassung, deren Folge einen beliebigen Dritten trifft und der vorsätzlichen Unterlassung, deren Folge eine dem Unterlassenden in irgendeiner Art verbundene Person trifft, dem sozialen Empfinden?
- 315 HAIDT und BARON sind dieser Frage nachgegangen. Sie unterscheiden – einer überbrachten Klassifizierung folgend – zwischen sozialen Beziehungen, die auf Hierarchie beruhen (vertikale Beziehungen) und sozialen Beziehungen, die auf Solidarität beruhen (horizontale Beziehungen).⁵⁵⁸ Sie stellen die Hypothese auf, dass Personen in einer Vorgesetzten-Rolle an den Folgen ihrer Handlungen *und* Unterlassungen gemessen werden, weshalb die Unterscheidung zwischen der Beurteilung von Handlungen und Unterlassungen bei ihnen weitgehend verschwindet. Umgekehrt sind Untergebene nicht in erster Linie für die Folgen ihrer Handlungen verantwortlich, sondern für die Einhaltung der ihnen erteilten Weisungen. Bei ihnen sollte die Unterscheidung zwischen Handlungen und Unterlassungen daher besonders ausgeprägt sein. Bei solidarischen Beziehungen, die auf Austausch und gegenseitiger Fürsorge beruhen, ist jede Person für die Beeinträchtigung der anderen Person durch ihre Handlungen und Unterlassungen verantwortlich, weshalb auch hier der Unterlassungseffekt weitgehend verschwinden sollte.⁵⁵⁹
- 316 Die Versuchspersonen wurden wiederum gebeten, den moralischen Wert einer Handlung oder Unterlassung, die zu negativen, aber eindeutig beabsichtigten, Folgen führte, auf einer

⁵⁵³ SPRANCA/MINSK/BARON, FN 545, 83.

⁵⁵⁴ SPRANCA/MINSK/BARON, FN 545, 103.

⁵⁵⁵ KORDES-DE VAAL, FN 528, 161 ff. Der Bindestrich im Namen ist übrigens am richtigen Ort.

⁵⁵⁶ KORDES-DE VAAL, FN 528, 169.

⁵⁵⁷ KORDES-DE VAAL, FN 528, 170.

⁵⁵⁸ JONATHAN HAIDT/JONATHAN BARON, Social Roles and the Moral Judgement of Acts and Omission, European Journal of Social Psychology 1996, 201-218, 202 f.

⁵⁵⁹ HAIDT/BARON, FN 558, 204.

Skala von -100 (sehr schlecht) bis +100 (sehr gut) zu bewerten. Neben Handlung/Unterlassung wurde die soziale Beziehung des Täters zum Geschädigten manipuliert: in einem Szenario waren die beiden „beste Freunde“, in einem zweiten „kannten sie sich aus der Nachbarschaft“ und in einem dritten waren sie sich noch nie begegnet. In allen drei Szenarien wurde die Handlung als schlechter bewertet als die vorsätzliche Unterlassung, die zu den gleichen Folgen führte. Der Unterschied in der Bewertung nahm aber ab, je enger die solidarische Beziehung von Opfer und Täter war: sie betrug im Schnitt 26,4 Punkte bei fehlender Beziehung, 19,7 Punkte bei mittlerer Beziehungsnähe und 15,5 Punkte bei enger Beziehung.⁵⁶⁰

- 317 Die Handlung des Untergebenen, die zur Verletzung des Vorgesetzten führte, wurde im Schnitt mit -41,2 Punkten bewertet, die vorsätzliche Unterlassung, die zur gleichen Verletzung führte, mit -30,4 Punkten. Die Handlung des Vorgesetzten in der gleichen Situation, die zur Verletzung des Untergebenen führte, wurde mit -51,9 bewertet, die vorsätzliche Unterlassung mit -44,4 Punkten.⁵⁶¹ Auch hier wurde in beiden Fällen die Handlung als moralisch schlechter als die vorsätzliche Unterlassung bewertet, aber der Unterschied zwischen der Bewertung von Handlung und Unterlassung ist beim Vorgesetzten geringer (7,5 Punkte) als beim Untergebenen (10,8 Punkte). Die Hypothese, dass soziale Beziehungen die Unterscheidung von Handlung und vorsätzlicher Unterlassung beeinflusst, wird also bestätigt.⁵⁶²

II. Rechtliche Beurteilung von Handlungen und vorsätzlichen Unterlassungen

- 318 Das Strafrecht unterscheidet traditionellerweise zwischen Handlungen und Unterlassungen. Der Satz „*maius delictum est in faciendo quam in omittendo*“ wird dem Glossatoren Accursius (1182-1221) zugeschrieben; ob er ihn als allgemeines Prinzip verstanden hat, bleibt freilich unklar.⁵⁶³ Die *Begründung* der Unterscheidung beschränkt sich allerdings bis in die neuste Zeit auf Evidenzargumente oder Appelle an das Rechtsgefühl.⁵⁶⁴ Das schweizerische Strafrecht sieht unter bestimmten, nachfolgend näher dargestellten, Bedingungen eine Bestrafung für vorsätzliche Unterlassungen vor. Jedes Vorsatzdelikt, das durch Begehung begangen werden kann, kann nach h.L. und Rechtsprechung auch durch Unterlassung begangen werden. Zudem gibt es seit 1990 den Straftatbestand der Unterlassung der Nothilfe, der nur durch vorsätzliche Unterlassung begangen werden kann.

⁵⁶⁰ HAIDT/BARON, FN 558, 207. Alle Unterschiede signifikant auf dem 5 %-Level.

⁵⁶¹ HAIDT/BARON, FN 558, 210. Alle Unterschiede signifikant auf dem 5 %-Level.

⁵⁶² Dies scheint auch ein kulturspezifisches Phänomen zu sein. HAIDT/BARON, FN 558, 217, erwähnen eine Studie von MILLER/BERSOFF/HARWOOD (1990) mit indischen Versuchspersonen, die keinen signifikanten Einfluss der sozialen Rolle auf die Bewertung vorsätzlicher Unterlassungen fand.

⁵⁶³ BIRNBACHER, FN 548, 119.

⁵⁶⁴ Kritisch BIRNBACHER, FN 548, 19, 118 ff.

A. Unterlassung der Nothilfe (Art. 128 StGB)

319 Die meisten Länder des angelsächsischen Rechtskreises kennen keine „*good samaritan*“ Gesetze, die zur Leistung von Nothilfe verpflichten.⁵⁶⁵ In Europa hat Portugal Mitte des 19. Jh. als erstes Land einen Straftatbestand der Unterlassung der Nothilfe geschaffen. Mittlerweile kennt eine Mehrheit der kontinentaleuropäischen Staaten solche Bestimmungen (in Deutschland § 323c dStGB).⁵⁶⁶ In der Schweiz wurde die Bestimmung zur Unterlassung der Nothilfe (Art. 128 Strafgesetzbuch, StGB) erst am 1. Januar 1990 Bestandteil des eidgenössischen Strafrechts. Eine strafrechtlich sanktionierte Verpflichtung, einem beliebigen Dritten, zu dem der Täter nicht in einer besonderen Beziehung steht, zu helfen, besteht nur in einem eng definierten Rahmen. Eine Bestrafung irgendeines Dritten – d.h. nicht desjenigen, der dem Geschädigten eine Verletzung zugefügt hat – wegen vorsätzlicher Unterlassung der Nothilfe setzt voraus, dass

- sich der Geschädigte in unmittelbarer Lebensgefahr befindet; und
- der Täter trotz Tatmacht, d.h. der Möglichkeit, dem Geschädigten zu helfen, die Hilfe unterlässt; und
- die Hilfeleistung dem Täter den Umständen nach zugemutet werden kann, ihn also beispielsweise nicht selber einer erheblichen Gefahr aussetzt.

320 Subjektiv setzt unterlassene Nothilfe Vorsatz voraus, d.h. insbesondere

- muss der Täter erkannt haben, dass sich der Geschädigte in unmittelbarer Lebensgefahr befindet; und
- er muss seine Tatmacht erkennen; und
- er muss die Zumutbarkeit der Hilfeleistung erkennen; und
- er muss trotz diesen Kenntnissen wissentlich und willentlich auf die Hilfeleistung verzichten.

321 Nicht vorausgesetzt wird hingegen, dass der Täter erkennt oder beabsichtigt, dass sich durch seine unterlassene Hilfeleistung die Lage des Geschädigten verschlechtert; Unterlassung der Nothilfe ist ein abstraktes Gefährdungsdelikt und umfasst keinen tatbestandsmäßigen Erfolg.⁵⁶⁷

322 In Anbetracht der restriktiven Voraussetzungen – in Deutschland ist die unmittelbare Lebensgefahr zum Beispiel nicht erforderlich – überrascht es kaum, dass Unterlassung der Nothilfe, begangen durch einen beliebigen Dritten, in der Praxis geringe Bedeutung hat. Das Bundesgericht hat sich seit dem Inkrafttreten der Bestimmung am 1. Januar 1990 in genau zwei publizierten Urteilen mit Art. 128 StGB befasst. In einem Fall entschied es, dass Art. 128 StGB einer kantonalen Bestimmung vorgeht, die die Unterlassung der Nothilfe auch in Fällen, in denen sich der Geschädigte nicht in unmittelbarer Lebensgefahr befindet, für strafbar erklärt. Es hob deshalb eine Verurteilung der Vorinstanz, die sich auf

⁵⁶⁵ FEINBERG, FN 549, 127.

⁵⁶⁶ FEINBERG, FN 549, 127.

⁵⁶⁷ GÜNTHER STRATENWERTH/GUIDO JENNY, Schweizerisches Strafrecht, Besonderer Teil I: Straftaten gegen Individualinteressen, 6. Aufl., Bern 2003, 91.

den kantonalen Straftatbestand gestützt hatte, auf (BGE 116 IV 19). In einem weiteren Fall befand das Bundesgericht, dass unmittelbare Lebensgefahr vorliegt, wenn eine Überdosis Heroin binnen weniger Stunden zum Tod führen wird (BGE 121 IV 18). Es handelt sich bei diesem Fall um den einzigen publizierten Entscheid, in dem das höchste Gericht eine Verurteilung wegen Art. 128 StGB bestätigte, ohne dass der Täter den Geschädigten selber verletzt hatte.⁵⁶⁸

B. Unechte Unterlassungsdelikte

323 Nach einhelliger Lehre und Rechtsprechung kann jedes vorsätzliche Erfolgsdelikt auch durch Unterlassung begangen werden, allerdings nur dann, wenn der Täter durch ausserstrafgesetzliche Gebote zur Verhinderung des Erfolgseintritts verpflichtet ist. Diese Pflicht kann sich nach dem Bundesgericht „aus gesetzlicher Vorschrift, aus Vertrag oder auch aus den Umständen ergeben“ (BGE 96 IV 155, 174).⁵⁶⁹ Der Täter des unechten Unterlassungsdelikts muss zum Geschädigten in einer so genannten Garantenstellung stehen; beim unechten Unterlassungsdelikt handelt es sich mithin um ein echtes Sonderdelikt.

324 In späteren Urteilen hat das Bundesgericht betont, dass die Garantenstellung nur durch eine „qualifizierte Rechtspflicht“ begründet werden kann (BGE 113 IV 68, 73):⁵⁷⁰

Ein unechtes Unterlassungsdelikt ist gegeben, wenn wenigstens die Herbeiführung des Erfolges durch Tun ausdrücklich mit Strafe bedroht wird, der Beschuldigte durch sein Tun den Erfolg tatsächlich hätte abwenden können und infolge seiner besonderen Rechtsstellung dazu auch so sehr verpflichtet war, dass die Unterlassung der Erfolgsherbeiführung durch aktives Handeln gleichwertig erscheint. Eine derartige Garantenstellung besteht insbesondere für den Täter, der kraft seiner besonderen Rechtsstellung das Gut vor der dieser drohenden Gefahr hätte schützen müssen. Mit der zitierten Rechtsprechung wird deutlich gemacht, dass nicht jede Rechtspflicht für die Begründung einer Garantenpflicht genügt, sondern nur eine qualifizierte.

325 Ein bloss moralisches Gebot zur Gefahrenabwendung genügt nicht, um die Garantenstellung zu begründen (BGE 108 IV 3, 6). Die Garantenpflichten lassen sich nach h.L., die vom Bundesgericht übernommen wird, nach Obhuts- und Überwachungspflichten unterscheiden (BGE 113 IV 68, 73):

Nach allgemeiner Ansicht lassen sich zwei Grundtypen von Garantenstellungen unterscheiden, nämlich einmal Obhutspflichten, d.h. Garantenstellungen zur Verteidigung bestimmter Rechtsgüter gegenüber unbestimmt vielen Gefahren, und zum anderen Überwachungspflichten, d.h. Garantenstellungen zur Überwachung bestimmter Gefahrenquellen zum Schutze unbestimmt vieler Rechtsgüter.

⁵⁶⁸ Im Urteil 6S.162/2000 vom 20. Dezember 2000 sprach das Bundesgericht einen Täter der Verletzung von Art. 128 StGB schuldig, der das Opfer während eines Raubes angeschossen und lebensgefährlich verletzt hatte (in Anbetracht dessen, dass der Täter zu insgesamt 14 Jahren Zuchthaus verurteilt wurde, fragt man sich, welche Strafschärfung wohl die Verurteilung wegen Art. 128 StGB ausgemacht hat).

⁵⁶⁹ JÖRG REHBERG, Strafrecht I, 5. Aufl. Zürich 1993, 182, kritisiert zu Recht, dass die Umschreibung „aus den Umständen“ in Anbetracht von Art. 1 StGB (keine Strafe ohne ausdrückliches Gesetz) viel zu unbestimmt ist.

⁵⁷⁰ Quellenhinweise des Bundesgerichts weggelassen.

- 326 Eine aussergesetzliche Begründung der Garantenstellung ergibt sich nach h.L. und Rechtsprechung in erster Linie aus der so genannten Ingerenz,⁵⁷¹ d.h. derjenige, der „durch sein Tun eine Gefahr geschaffen oder eine solche vergrössert hat [ist] deshalb gehalten [...], dafür zu sorgen, dass die Gefahr zu keiner Verletzung fremder Rechtsgüter führt“ (BGE 108 IV 3, 5).⁵⁷²
- 327 Die rechtliche Regelung der Strafbarkeit deckt sich daher mit dem sozialen Empfinden und wendet sich gegen philosophische Konzepte, die Handlung und vorsätzliche Unterlassung gleichsetzen (insbesondere auch bei der unterschiedlichen Behandlung von aktiver und passiver Sterbehilfe, die für den Ethiker PETER SINGER nicht zu rechtfertigen ist, nach schweizerischem Strafrecht aber anerkannt wird).⁵⁷³
- 328 Die Übereinstimmung zwischen sozialem Empfinden und dem Recht der unechten Unterlassungsdelikte wird noch frappierender, wenn man einzelne höchstrichterliche Urteile betrachtet. Die Begründung von Obhuts- und Schutzpflichten aus allgemeinen Rechtspflichten hat oft etwas Aleatorisches.⁵⁷⁴ In BGE 96 IV 155, 174 f. argumentiert das Bundesgericht, die Pflicht von Bührle, gegen unbewilligte Lieferungen von Waffen nach Südafrika durch Mitarbeiter seines Konzerns einzuschreiten, ergebe sich vielleicht nicht aus positivem Recht, aber aus seiner Stellung als „oberster Leiter“, „nach aussen in Erscheinung tretende[m] Inhaber“ und seiner „beherrschenden Rolle“ „als Haupt der [Waffenfabrik]“. Mit anderen Worten wird, wie dies dem sozialen Empfinden entspricht, eine Person in einer vorgesetzten Position an den Folgen ihrer Entscheidungen gemessen, handle es sich um Handlungen oder vorsätzliche⁵⁷⁵ Unterlassungen.
- 329 Andererseits hat das Bundesgericht in BGE 113 IV 68 entschieden, dass die arbeitsvertragliche Treuepflicht des Arbeitnehmers (Art. 321a Abs. 1 Obligationenrecht, OR) keine Garantenstellung des Arbeitnehmers für das Vermögen des Arbeitgebers begründet. Nur wer eine leitende Stellung innehat, hat eine Garantenpflicht für das Vermögen des Arbeitgebers, sofern die schädigenden Handlungen in seinem Zuständigkeitsbereich vorgenommen wurden (BGE 113 IV 68, 75). Gleich argumentiert auch REHBINDER, der nur den leitenden Angestellten als verpflichtet ansieht, gegen Schädigungen des Arbeitgebers durch andere Arbeitnehmer einzuschreiten, während die Pflicht, die Interessen des Arbeitgebers nicht aktiv zu schädigen, für jeden Arbeitnehmer ungeachtet seiner Stellung gilt.⁵⁷⁶ Auch hier deckt sich die Rechtsprechung mit dem sozialen Empfinden, das den

⁵⁷¹ Zum Begriff GÜNTER STRATENWERTH, Schweizerisches Strafrecht, Allgemeiner Teil I: Die Straftat, 2. Aufl., Bern 1996, 411 ff.

⁵⁷² In Anbetracht des Rückschaufehlers ist es generell sehr problematisch, aus einer Gefahrenlage auf das Bestehen von Rechtspflichten zu schliessen, siehe dazu hinten, S. 227 f. Kritisch auch STRATENWERTH, FN 571, 412 ff.

⁵⁷³ Zur strafrechtlichen Beurteilung der Sterbehilfe in der Schweiz siehe ANDREAS DONATSCH, Die strafrechtlichen Grenzen der Sterbehilfe, in: MATTHIAS METTNER (Hrsg.), Wie menschenwürdig sterben?: Zur Debatte um die Sterbehilfe und die Praxis der Sterbebegleitung, Zürich 2000, 121-137.

⁵⁷⁴ STRATENWERTH, FN 571, 408: „Die Voraussetzungen, unter denen eine Garantenpflicht anzunehmen ist, sind, wie schon angedeutet, sehr zweifelhaft und umstritten“.

⁵⁷⁵ Bührle wusste gemäss Sachverhalt spätestens seit dem 1. oder 2. Juli 1965 mit Sicherheit von den verbotenen Lieferungen, BGE 96 IV 155, 176.

⁵⁷⁶ MANFRED REHBINDER, Einleitung und Kommentar zu den Art. 319-330a OR, Bern 1985, Art. 321c N 6 und 9.

Untergebenen nicht für die Folgen seiner vorsätzlichen Unterlassungen verantwortlich macht.

- 330 Rechtsprechung und soziales Empfinden stehen im Widerspruch zueinander, wenn es um die Beurteilung der vorsätzlichen Unterlassung einer Person geht, die in einer rechtlich sanktionierten solidarischen Beziehung zu einer anderen Person steht – Paradebeispiel ist natürlich die Ehe, wo Art. 159 Abs. 3 Zivilgesetzbuch (ZGB) eindeutig gegenseitige Beistandspflichten der Eheleute statuiert. Wie die Experimente von HAIDT/BARON zeigen, beurteilen Menschen die moralische Vorwerfbarkeit von Handlung und vorsätzlicher Unterlassung in solchen Beziehungen immer noch unterschiedlich, wenn auch weniger stark als ohne solidarische Beziehung, während das Recht verlangt, dass sie gleich zu beurteilen sind.
- 331 Die Begehung durch vorsätzliche Unterlassung ist kein Strafmilderungsgrund nach Art. 64 StGB (anders in Deutschland, wo § 13 Abs. 2 dStGB die fakultative Strafmilderung für unechte Unterlassungsdelikte ausdrücklich vorsieht). Die Lehre erwähnt auch nicht, dass die Begehung durch Unterlassung als strafmindernd im Rahmen von Art. 63 StGB zu berücksichtigen ist, Faktoren wie die „kriminelle Energie“ des Täters und die „Intensität des verbrecherischen Willens“⁵⁷⁷ werden allerdings auch durch die Unterscheidung von Handlung und vorsätzlicher Unterlassung beeinflusst. Die psychologische Forschung zum Unterlassungseffekt legt es nahe, dass Richterinnen und Richter dennoch einen Unterschied beim Verschulden machen werden.

III. Eigene Studie zum Unterlassungseffekt

A. Frage zum Unterlassungseffekt

- 332 Den Richterinnen und Richtern der Kantone Basel-Landschaft, Bern und Graubünden wurde folgender Sachverhalt zur Beurteilung vorgelegt (2004):

Das Ehepaar E. ist seit vielen Jahren verheiratet. Die Ehe ist nicht mehr glücklich und der Ehemann hat eine Geliebte. Er würde sich am liebsten scheiden lassen, verzichtet aber aus finanziellen Gründen darauf.

Die Ehefrau leidet an einer Erdnuss-Allergie. Diese kann beim Verzehr von Erdnüssen zu einem allergischen Schock und zum Tod führen. Ihr Ehemann weiss dies.

Das Stammlokal des Ehemannes macht eine hausgemachte Salatsauce („House-Dressing“), die einen geringen Anteil an pürierten Erdnüssen enthält. Der Ehemann weiss von den Erdnüssen im House-Dressing.

Auf Wunsch der Ehefrau geht das Ehepaar gemeinsam in das Stammlokal des Ehemannes.

Als die Ehefrau einen Salat bestellt, fragt der Kellner, ob sie italienische oder französische Sauce oder das House-Dressing möchte. Die Ehefrau weiss nicht, dass das House-Dressing Spuren von Erdnüssen enthält und erkundigt sich auch nicht danach.

- 333 In einem zweifaktoriellen Versuchsplan wurden die Faktoren „Handlung versus Unterlassung“ und „Erfolg versus Versuch“ variiert. In der Version „Unterlassung“ lautete der Sachverhalt:

⁵⁷⁷ STEFAN TRECHSEL, Schweizerisches Strafgesetzbuch: Kurzkommentar, Zürich 1989, vor Art. 63 N 12 f.

Besonderer Teil

Die Frau bestellt das House-Dressing. Der Mann warnt sie nicht. Der Mann weiss, dass sich seine Frau für eine andere Sauce entschieden hätte, wenn er sie gewarnt hätte.

334 In der Version „Handlung“ lautete der Sachverhalt stattdessen:

Bevor sich die Frau entscheiden kann, empfiehlt ihr Mann das House-Dressing. Die Frau entscheidet sich dann auch für das House-Dressing. Hätte der Mann nichts gesagt, so hätte seine Frau die italienische Sauce bestellt.

335 In der Version „Erfolg“ erleidet die Frau einen allergischen Schock, an dem sie stirbt. In der Version „Versuch“ erleidet sie ebenfalls einen allergischen Schock, der aber sehr schwach ausfällt und ausser einem kurzzeitigen Unwohlsein keine weiteren Folgen hat.

336 Die Richterinnen und Richter wurden zuerst gefragt, wie sie das Verschulden des Mannes auf einer siebenstufigen Skala (1 = geringes Verschulden, 7 = schweres Verschulden) einschätzen. Zweitens wurden sie nach der ihrer Ansicht nach angemessenen Strafe gefragt (offene Fragestellung).

337 Die Hypothese ist natürlich, dass die Richter das Verschulden bei der Unterlassung geringer als bei der Handlung bewerten und eine tiefere Strafe aussprechen. Erwartet wurde ebenfalls, dass Verschulden und Strafe im Erfolgsfall (Tod der Frau) schwerer ausfällt als beim vollendeten Versuch.

B. Rechtliche Beurteilung der Sachverhaltsvarianten

338 In der Version „Handlung, Erfolg“ des Sachverhalts macht sich der Mann der vorsätzlichen Tötung im Sinne von Art. 111 StGB, evtl. des Mordes im Sinne von Art. 112 StGB schuldig.

339 Art. 111 StGB ist ein Erfolgsdelikt, Erfolg ist die Tötung eines anderen Menschen. Der Mann unternimmt alles, was seiner Meinung nach notwendig ist, um den tatbestandsmässigen Erfolg herbeizuführen, indem er der Frau die erdnusshaltige Salatsauce empfiehlt. Dass er mit Wissen und Willen handelt, kann aufgrund der Sachverhaltsschilderung kaum bezweifelt werden. Er weiss um die Erdnussallergie seiner Frau und deren potenziell tödlichen Folgen bei Genuss von Erdnüssen. Indem er seiner Frau trotz dieses Wissens die Erdnusssauce empfiehlt, nimmt er die Tötung der Frau mindestens billigend in Kauf und handelt somit eventualvorsätzlich.

340 Fragen kann man sich, ob der Mann gar besonders skrupellos oder verwerflich im Sinne des qualifizierten Tatbestandes von Art. 112 StGB (Mord) handelt. Der Sachverhalt legt nahe, dass er die Tötung aus finanziellen Überlegungen, um die Scheidungskosten zu sparen, begeht. In BGE 77 IV 64 hat das Bundesgericht die Tötung eines Ehemannes, der einer Heirat des Geliebten entgegenstand, durch die Ehefrau als Mord qualifiziert; vgl. zur Umschreibung des Mordes generell BGE 118 IV 122, 126, gemäss dem sich ein Mörder in der Regel durch kaltblütigen Egoismus „auszeichnet“. Meines Erachtens liegt es daher nahe, hier einen Mord im Sinne von Art. 112 StGB anzunehmen; aber da die Abgrenzung zwischen vorsätzlicher Tötung und Mord letztlich immer ein Ermessensentscheid bleibt und für das Experiment keine Rolle spielt, möchte ich die Frage nicht weiter vertiefen.

- 341 In der Version „Handlung, Versuch“ liegt vollendeter Versuch der vorsätzlichen Tötung resp. des Mordes sowie eine vollendete Tätlichkeit vor (Art. 111 i. V. m. Art. 22 Abs. 1, Art. 126 StGB). Der Mann unternimmt alles, was seiner Meinung nach zur Herbeiführung des tatbestandsmässigen Erfolgs notwendig ist, aber der Erfolg tritt aufgrund von Umständen, die ausserhalb seiner Herrschaft liegen, nicht ein. Art. 22 StGB sieht für Fälle des vollendeten Versuchs eine fakultative Strafmilderung vor („so kann der Täter milder bestraft werden“). Das Bundesgericht legt Art. 22 StGB aber dahingehend aus, dass die Strafmilderung beim Versuch obligatorisch ist (BGE 121 IV 49, 55). Gemäss Bundesgericht hängt „das Mass der zulässigen Reduktion der Strafe beim vollendeten Versuch unter anderem von der Nähe des tatbestandsmässigen Erfolgs und den tatsächlichen Folgen der Tat ab. Die Reduktion der Strafe wird mit andern Worten umso geringer sein, je näher der tatbestandsmässige Erfolg und je schwerwiegender die tatsächliche Folge der Tat war“ (BGE 121 IV 49, 54). Tatsächlich verursacht die Handlung des Mannes ein „kurzzeitiges Unwohlsein“ der Frau, das wohl noch in den Bereich der Tätlichkeit im Sinne von Art. 126 StGB fällt – die Abgrenzung der Tätlichkeit zur einfachen Körperverletzung ist bekanntlich nicht einfach, und dem Richter kommt ein erhebliches Ermessen zu.
- 342 In der Version „Unterlassung, Erfolg“ macht sich der Mann der vorsätzlichen Tötung, evtl. des Mordes, begangen durch Unterlassung, schuldig (Art. 111, evtl. 112 StGB). Die Garantenstellung des Ehemannes ergibt sich aus Art. 159 Abs. 3 ZGB; dass Art. 159 ZGB eine Garantenstellung begründet, wird nicht bezweifelt.⁵⁷⁸ Der Mann nimmt die Tötung seiner Frau durch Unterlassung zumindest eventualvorsätzlich in Kauf. Die Tatmacht, den Eintritt des Erfolgs zu verhindern, hat er ebenfalls, er bräuchte seine Frau nur zu warnen.
- 343 In der Version „Unterlassung, Versuch“ macht sich der Mann des vollendeten Versuchs der vorsätzlichen Tötung, evtl. des Mordes, begangen durch Unterlassung, sowie der Tätlichkeit, ebenfalls begangen durch Unterlassung, schuldig (Art. 111, evtl. 112, i. V. m. Art. 22 Abs. 1, Art. 126 StGB). Der Mann unterlässt alles, was seiner Meinung nach notwendig ist, den Tod seiner Frau herbeizuführen. Dass die Frau nicht stirbt, ist auf einen glücklichen Umstand zurückzuführen, der ausserhalb der Einflussphäre des Ehemannes liegt. Dass der vollendete Versuch eines unechten Unterlassungsdelikts geringe praktische Bedeutung hat, ist klar, dogmatisch ist er zweifellos möglich.⁵⁷⁹

C. Resultate

- 344 Von den 234 Richterinnen und Richter, die den Fragebogen zurückschickten, haben 226 (96,6 %) die Frage nach dem Verschulden beantwortet. Die Frage nach der angemessenen Strafe haben 63 (28 %) der Richterinnen und Richter, die die Frage nach dem Verschulden beantwortet haben, nicht beantwortet, und zwar zehn der Gruppe „Handlung, Erfolg“, 15 der Gruppe „Handlung, Versuch“, zwölf der Gruppe „Unterlassung, Erfolg“ und neun der Gruppe „Unterlassung, Versuch“. Acht Richter in der Gruppe „Unterlassung, Versuch“

⁵⁷⁸ STRATENWERTH, FN 571, 409; STEFAN TRECHSEL/PETER NOLL, Schweizerisches Strafrecht, Allgemeiner Teil I, 6. Aufl., Zürich 2004, 251, schränken die Garantenstellung auf tatsächlich zusammenlebende Ehegatten ein.

⁵⁷⁹ TRECHSEL/NOLL, FN 578, 262.

sahen eine Busse als angemessene Strafe, drei Richter in der Gruppe „Handlung, Versuch“ ebenfalls. Kein Richter in der Gruppe „Unterlassung, Erfolg“ oder „Handlung, Erfolg“ erachtete eine Busse als angemessen.⁵⁸⁰ Die Antworten der Richterinnen und Richter, die eine Busse aussprachen, wurden für die folgende Auswertung eliminiert.

Die Resultate sind in der Tabelle 9 von oben nach unten in der Reihenfolge der Beurteilung der Schwere des Verschuldens dargestellt. Die Version des Sachverhalts, bei der das Verschulden als am schwerwiegendsten beurteilt wurde, ist zuoberst. Die kleinere Anzahl verwertbarer Antworten bei den Freiheitsstrafen erklärt sich durch die oben erwähnten fehlenden Antworten, resp. auf Busse erkannten Strafen.

Gruppe		Verschulden	Freiheitsstrafe (in Monaten)
Handlung, Erfolg	Durchschnitt	5,84	91,35
	N	59	49
	Standardabw.	1,29	69,96
Unterlassung, Erfolg	Durchschnitt	5,77	61,27
	N	59	45
	Standardabw.	1,27	54,90
Handlung, Versuch	Durchschnitt	4,51	14,17
	N	53	30
	Standardabw.	1,69	18,58
Unterlassung, Versuch	Durchschnitt	4,35	13,23
	N	55	39
	Standardabw.	1,87	19,31
Total	Durchschnitt	5,15	50,15
	N	226	163
	Standardabw.	1,68	59,44

Tabelle 9: Resultate Unterlassungseffekt

- 345 Auf den ersten Blick fällt auf, dass die Folge einen sehr viel grösseren Einfluss auf das Verschulden und das Strafmass hat als die Unterscheidung zwischen Handlung und Unterlassung. Bei Erfolg – sprich Tod der Ehefrau – wird das Verschulden im Schnitt mit 5,81 Punkten (N = 118) bewertet und eine Freiheitsstrafe von 76,95 Monaten (N = 94) ausgesprochen, während beim Versuch (kurzzeitiges Unwohlsein) das Verschulden mit 4,43 Punkten (N = 108) bewertet und eine Freiheitsstrafe von 13,64 Monaten (N = 69) ausgesprochen wird. Der Einfluss des Ausgangs auf das Verschulden⁵⁸¹ wie auch auf das Strafmass⁵⁸² ist signifikant.

⁵⁸⁰ Der Unterschied ist statistisch signifikant, $\chi^2 = 12,790$, $p < 0,001$.

⁵⁸¹ Mann-Whitney U = 3'391,5, Wilcoxon W = 9'277,5, $p \leq 0,001$.

⁵⁸² Mann-Whitney U = 897,0, Wilcoxon W = 3'312,0, $p \leq 0,001$.

- 346 Der Einfluss der Variable Handlung/Unterlassung auf Verschulden und Strafmass ist geringer. Das Verschulden bei den Handlungen wurde im Schnitt mit 5,21 Punkten ($N = 112$) bewertet und eine Strafe von 38,97 Monaten ($N = 84$) ausgesprochen. Bei den Unterlassungen wurde das Verschulden im Schnitt mit 5,09 ($N = 114$) bewertet und eine Strafe von 62,04 Monaten ($N = 79$) ausgesprochen. Der Trend des Einflusses zeigt also sowohl beim Verschulden wie auch bei der Unterlassung in die erwartete Richtung, ist aber nur beim Strafmass statistisch signifikant.⁵⁸³
- 347 Betrachtet man die Sachverhalte mit Erfolg und diejenigen, bei denen es beim Versuch blieb, getrennt, zeigt sich folgendes Bild: der Unterlassungseffekt zeigt bei den Sachverhalten mit harmlosem Ausgang weder auf das Verschulden noch auf das Strafmass einen signifikanten Einfluss,⁵⁸⁴ hingegen wirkt er sich auf das Strafmass, nicht jedoch auf das Verschulden, bei den Sachverhalten mit tödlichem Ausgang signifikant aus.⁵⁸⁵

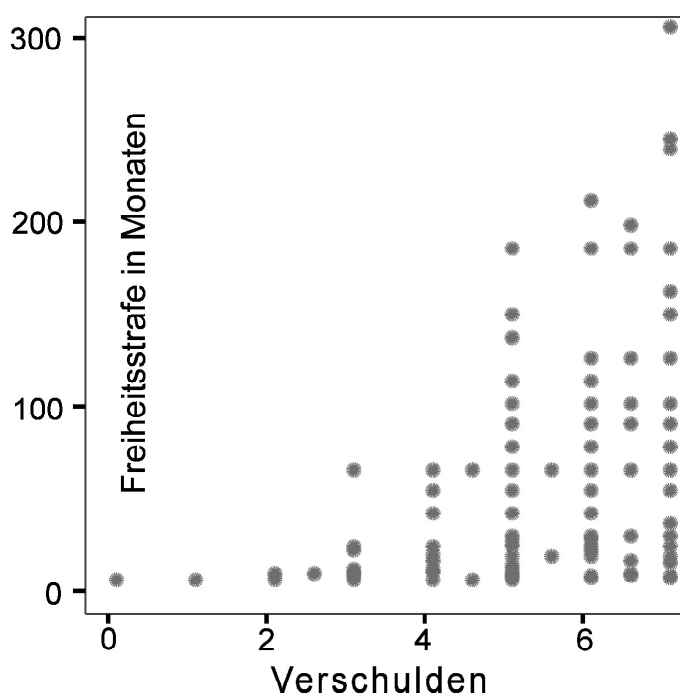


Abbildung 9: Streudiagramm von Verschulden x Strafmass

- 348 Die Korrelation zwischen Verschulden und Strafmass ist in der Abbildung 9 grafisch dargestellt. Der Rangkorrelationskoeffizient (Spearman's Rho) zwischen Verschulden und Strafmass beträgt 0,62 und ist signifikant ($p \leq 0,01$).

⁵⁸³ Beim Verschulden: Mann-Whitney $U = 6'147,0$, Wilcoxon $W = 12'702,0$, $p = 0,624$; beim Strafmass: Mann-Whitney $U = 2'675,5$, Wilcoxon $W = 6'245,5$, $p = 0,033$.

⁵⁸⁴ Beim Verschulden: Mann-Whitney $U = 1403,0$, Wilcoxon $W = 2943,0$, $p = 0,735$; beim Strafmass: Mann-Whitney $U = 540,0$, Wilcoxon $W = 1320,0$, $p = 0,583$.

⁵⁸⁵ Beim Verschulden: Mann-Whitney $U = 1636,5$, Wilcoxon $W = 3406,5$, $p = 0,563$; beim Strafmass: Mann-Whitney $U = 821,5$, Wilcoxon $W = 1856,5$, $p = 0,033$.

D. Diskussion

- 349 Als erstes möchte ich klar festhalten, dass die obigen Resultate nicht einen wie auch immer gearteten „Fehler“ oder eine „Irrationalität“ der Richterinnen und Richter belegen. Es gibt schlicht keinen exakten normativen Standard, an dem sich Verschulden oder Strafzumessung messen liessen, folglich kann man auch nicht von einer Abweichung oder einem Fehler sprechen.⁵⁸⁶ Die Resultate sind aber hochinteressant, weil sie den Einfluss des sozialen Empfindens auf das Strafmass zeigen und gewisse gesetzliche Unterscheidungen (wie diejenige zwischen aktiver und passiver Sterbehilfe) erklären können.
- 350 Die Richterinnen und Richter haben den erfolgreichen Täter im Schnitt zu 77 Monaten, also rund sechseinhalb Jahren, Freiheitsstrafe verurteilt, während der vollendete Versuch im Schnitt mit 13,6 Monaten oder etwas mehr als einem Jahr bestraft wurde. Das Strafmass wurde viel stärker vom Ausgang als von der Unterscheidung Handlung/Unterlassung bestimmt, was sich durch zwei Thesen erklären lässt: entweder unterliegen die Richterinnen und Richter einem Ausgangseffekt,⁵⁸⁷ d.h. sie beurteilen eine Tat in erster Linie nach ihren Folgen, oder sie beurteilen die Strafwürdigkeit einer Tat in erster Linie nach dem Ausmass der von ihr bewirkten Störung des Rechtsfriedens.
- 351 Nach dem reinen Willensstrafrecht bemisst sich die Strafe ausschliesslich nach dem „bösen“ Willen und der Gefährlichkeit des Täters, es spielt somit keine Rolle, ob sich der beabsichtigte Erfolg verwirklicht hat oder nicht. Vollendete Tat und vollendeter Versuch werden nach dieser Konzeption, die derjenigen des französischen Code pénal von 1810 entspricht, gleich bestraft.⁵⁸⁸
- 352 Stellt man stärker auf die Bewahrung des Rechtsfriedens ab, so ist der Versuch aus objektiven Gründen strafbar, weil auch ein Versuch den Rechtsfrieden stört; er wirkt gefährlich und erregt Furcht.⁵⁸⁹ Aber er stört den Rechtsfrieden natürlich viel weniger als die erfolgreiche Straftat, in den Worten von TRECHSEL/NOLL:⁵⁹⁰
- Aber es macht denn doch einen sehr erheblichen Unterschied, ob auf mich geschossen wurde oder ob ich erschossen worden bin. Die Empörung über den Versuch wird gemildert durch die Erleichterung, dass es dabei sein Bewenden hatte.
- 353 Unter diesem Gesichtspunkt ist klar, dass der Versuch weniger hart bestraft wird als die vollendete Tat. TRECHSEL/NOLL müssen zugeben, dass sich aus subjektiver Sicht die geringere Bestrafung des Versuchs kaum rechtfertigen lässt; unter Hinweis auf FREUD machen sie geltend, dass unbewusste rechtstreue Strebungen des Täters dazu beigetragen haben mögen, dass der Erfolg ausgeblieben ist.⁵⁹¹
- 354 Die Unterscheidung zwischen Handlung und Unterlassung hat sich wie erwähnt auf das *Verschulden* nicht signifikant ausgewirkt. Der Grund dürfte darin zu suchen sein, dass die

⁵⁸⁶ Zu den Schwierigkeiten, die Strafzumessung zu objektivieren siehe z.B. TRECHSEL, FN 577, vor Art. 63.

⁵⁸⁷ Mehr dazu hinten, S. 210 f.

⁵⁸⁸ BERNARD BOULOC, Droit pénal général, 15. Aufl., Paris 1994, 207.

⁵⁸⁹ TRECHSEL/NOLL, FN 578, 177.

⁵⁹⁰ TRECHSEL/NOLL, FN 578, 177.

⁵⁹¹ TRECHSEL/NOLL, FN 578, 177.

verwendete Antwort-Skala zu wenig sensitiv war. Mit einer feineren Unterteilung – z.B. einer Skala von 1-100, wie sie RITOV/BARON verwendeten – hätte sich möglicherweise ein signifikanter Unterschied erkennen lassen. Darauf deutet hin, dass der Unterschied bei der (offenen) Frage nach dem Strafmass signifikant war. Auf der siebenstufigen Verschuldens-Skala gruppieren sich die Werte fast alle im oberen Drittel. 119 (52 %) aller Richterinnen und Richter bewerteten das Verschulden mit sechs oder mehr Punkten, aber nur 16 (7 %) mit zwei oder weniger Punkten. Die Skalenbegrenzung hat daher wahrscheinlich zur Reduktion der Varianz und damit zur fehlenden Signifikanz geführt („*ceiling effect*“).

- 355 Der Einfluss der Variable Handlung/Unterlassung auf das Strafmass ist bei erfolgreicher Tat sowohl erheblich als auch statistisch signifikant. Im „Handlungs“ Szenario wurde der Täter im Schnitt zu 91,35 Monaten, oder rund siebeneinhalb Jahren, Gefängnis verurteilt, im „Unterlassungs“ Szenario zu 61,27 Monaten (rund fünf Jahren). Die Strafe für die aktive Begehung fiel somit, von unten betrachtet, rund 50 % höher aus als die Strafe für die Begehung durch vorsätzliche Unterlassung des gleichen Delikts mit den gleichen Folgen. Die Resultate der Experimente von SPRANCA et al. konnten repliziert werden (der Leser wird unschwer erkannt haben, dass der hier verwendete Sachverhalt eine Abwandlung des „Ivan Lendl“ Szenarios von SPRANCA et al. ist).⁵⁹²
- 356 Hingegen ist der Unterschied in der Bestrafung beim vollendeten Versuch, begangen durch Handlung, zum vollendeten Versuch begangen durch Unterlassung gering und statistisch nicht signifikant. Die versuchte Begehung durch Unterlassung wurde im Schnitt mit 13,23 Monaten, die versuchte Begehung durch Handlung mit 14,17 Monaten bestraft. Dieses Resultat deckt sich mit der psychologischen Forschung, die feststellt, dass der Unterlassungseffekt bei negativem Ausgang stärker ist.⁵⁹³

1. Unterlassungseffekt und Sterbehilfe

- 357 Die unterschiedliche Bestrafung von Handlung und vorsätzlicher Unterlassung, die zum gleichen Resultat führen, entspricht offenbar einem menschlichen Bedürfnis. In der Schweiz spiegelt es sich beispielsweise darin, dass aktive Sterbehilfe, auch wenn sie dem ausdrücklichen Wunsch des Sterbenden entspricht, strafbar ist (Tötung auf Verlangen, Art. 114 StGB), während die passive Sterbehilfe, d.h. der vorsätzliche Verzicht auf die Anwendung lebenserhaltender oder –verlängernder Massnahmen, und die so genannte indirekte Sterbehilfe, d.h. das Verabreichen schmerzlindernder Medikamente, deren Nebenwirkungen schliesslich zum Tod führen, straflos bleibt.⁵⁹⁴
- 358 Darüber, ob die Unterscheidung zwischen aktiver und passiver Sterbehilfe angemessen ist, wird bekanntlich eine intensive Debatte geführt. Einen Überblick über die verschiedenen politischen Vorstösse und die Rechtslage in anderen Ländern bietet der Bericht der Ar-

⁵⁹² SPRANCA/MINSK/BARON, FN 545.

⁵⁹³ JONATHAN BARON/ILANA RITOV, Reference points and omission bias, *Organizational Behavior and Human Decision Making Processes* 1994, 475-498, 487.

⁵⁹⁴ DONATSCH, FN 573, 121 ff.

beitsgruppe Sterbehilfe an das EJPD vom März 1999.⁵⁹⁵ Die Mehrheit der Arbeitsgruppe kommt zum Schluss, dass unter sehr restriktiven Bedingungen (Tod steht kurz bevor, der Sterbende wird von unerträglichen Schmerzen befreit) direkte aktive Sterbehilfe straflos (aber weiterhin rechtswidrig) sein soll.

- 359 Gegen die Straflosigkeit der direkten aktiven Sterbehilfe haben sich vor allem christliche Kreise gewandt.⁵⁹⁶ Der Bundesrat empfahl im Bericht zum Postulat Ruffy, die passive und indirekte aktive Sterbehilfe zu regeln, aber die direkte aktive Sterbehilfe, entgegen dem Vorschlag der Mehrheit der Arbeitsgruppe, auch in extremen Fällen nicht zu erlauben. Dies „käme einer Lockerung des Fremdtötungsverbotes gleich und würde somit ein Tabu brechen, das in unserer christlichen Kultur tief verankert ist“.⁵⁹⁷ Die vorgeschlagene Regelung der passiven und indirekten Sterbehilfe ist bislang nicht erfolgt, nachdem die mit der Ausarbeitung der Grundlagen für die Regelung betraute Kommission im Februar 2004 mangels finanzieller Mittel ihres Auftrags entbunden wurde.⁵⁹⁸
- 360 Die Unterscheidung zwischen Handlung und Unterlassung lässt sich logisch-rational kaum begründen, wie selbst ein katholischer Theologe eingestehen muss.⁵⁹⁹ Für die Beibehaltung der herkömmlichen Unterscheidung spricht aber – neben moraltheologischen Argumenten und der Angst vor der „*slippery slope*“, die ich in Anbetracht des steigenden Kostendrucks in der Geriatriepflege nicht für unberechtigt halte⁶⁰⁰ – die offenbar tief verwurzelte Vorliebe des Menschen für Unterlassungen gegenüber Handlungen, die negative Konsequenzen haben und die ebenfalls experimentell nachweisbare Bevorzugung von indirektem gegenüber direktem Schaden.⁶⁰¹ Ob diese Unterscheidung in unserer christlichen Tradition begründet ist, muss einstweilen offen bleiben – alle hier referierten Experimente wurden in Ländern semitisch-christlicher Prägung durchgeführt, während eine indische Studie offenbar keinen Unterschied zwischen Handlung und Unterlassung finden konnte.⁶⁰²
- 361 Die empfundenen Unterschiede müssen ernst genommen werden. Strafgesetze sollen (auch) das unter Strafe stellen, was als strafwürdig empfunden wird. Den Leuten, die für die Beibehaltung der Unterscheidung eintreten vorzuwerfen, sie versuchten heldenhaft, einen prinzipiellen Unterschied zu finden, wo keiner besteht, ist zu einfach.⁶⁰³ Man kann auch für die Beibehaltung der Unterscheidung eintreten, ohne dass man einen prinzipiellen

⁵⁹⁵ Arbeitsgruppe Sterbehilfe, Bericht an das Eidgenössische Justiz- und Polizeidepartement, März 1999 (erhältlich unter www.ofj.admin.ch/themen/stgb-sterbehilfe/b-bericht-d.pdf, besucht am 3. Februar 2005).

⁵⁹⁶ Siehe Bericht des Bundesrates zum Postulat Ruffy, Sterbehilfe, Ergänzung des Strafgesetzbuchs, vom Juli 2000, 9 ff. (erhältlich unter www.ofj.admin.ch/themen/stgb-sterbehilfe/ber-ruffy-d.pdf, besucht am 3. Februar 2005).

⁵⁹⁷ Bericht des Bundesrates zum Postulat Ruffy, FN 596, 15.

⁵⁹⁸ www.ofj.admin.ch/themen/stgb-sterbehilfe/intro-d.htm (besucht am 3. Februar 2005).

⁵⁹⁹ ZIMMERMANN-ACKLIN, FN 548, 276.

⁶⁰⁰ BIRNBACHER, FN 548, 358, misst dem „Dammbruchargument“ ebenfalls die grösste Bedeutung zu, weist aber zu Recht darauf hin, dass auch bei der passiven Sterbehilfe, die allgemein als zulässig erachtet wird, die Gefahr besteht, dass dem Patienten insinuiert wird, es sei langsam Zeit zu gehen.

⁶⁰¹ EDWARD ROYZMAN/JONATHAN BARON, The Preference for Indirect Harm, Social Justice Research 2002, 165-184.

⁶⁰² HAIDT/BARON, FN 558, 217. Es wäre äusserst spannend, die Experimente von RITOV/BARON in Ländern nicht christlicher Tradition zu wiederholen.

⁶⁰³ So aber SUNSTEIN, FN 547, 1585.

Unterschied *findet*, sondern weil man einen Unterschied *empfindet*, und möchte, dass das geltende Strafrecht diesen Unterschied respektiert und reflektiert.

§ 11 Denken in Repräsentativität

I. Ein Beispiel zum Einstieg

362 Für PIERRE-SIMON LAPLACE war Wahrscheinlichkeitstheorie „im Grunde nur der Berechnung unterworfenen gesunden Menschenverstand“;⁶⁰⁴ etwas, das vernünftige Menschen instinktiv wissen, ohne genau erklären zu können, wie. LAPLACE hat hier wahrscheinlich etwas vorschnell von sich auf andere geschlossen. Viele Menschen empfinden die aus der Wahrscheinlichkeitstheorie folgenden Schlüsse als kontraintuitiv.⁶⁰⁵

363 Die folgende Frage, in der psychologischen Literatur als „Taxi-Problem“ bekannt, veranschaulicht die Schwierigkeiten des korrekten Denkens in Wahrscheinlichkeiten:

Zwei Taxigesellschaften sind in einer Stadt tätig. Die Taxis der Gesellschaft A sind grün, die der Gesellschaft B blau. Die Gesellschaft A stellt 15 % der Taxis, die Gesellschaft B die verbleibenden 85 %. Eines Nachts kommt es zu einem Unfall mit Fahrerflucht. Das fliehende Auto war ein Taxi. Ein Zeuge sagt aus, es habe sich um ein grünes Taxi gehandelt.

Das Gericht lässt den Zeugen auf seine Fähigkeit untersuchen, grüne und blaue Taxis unter nächtlichen Sichtbedingungen zu unterscheiden. Das Untersuchungsergebnis ist: In 80 % der Fälle identifiziert der Zeuge die Farbe zutreffend, in 20 % der Fälle irrt er sich.

Wie hoch ist die Wahrscheinlichkeit, dass es sich beim fliehenden Taxi um ein Taxi der Gesellschaft A gehandelt hat?

364 Die Mehrheit der befragten Personen antwortet, dass die Wahrscheinlichkeit 80 % beträgt.⁶⁰⁶ Die richtige Lösung⁶⁰⁷ ergibt sich aus dem so genannten Bayes-Theorem, benannt nach dem presbyterianischen Pfarrer und Mathematiker THOMAS BAYES (1702-1761).⁶⁰⁸

⁶⁰⁴ PIERRE-SIMON LAPLACE, Philosophischer Versuch über die Wahrscheinlichkeit, Leipzig 1932, 176.

⁶⁰⁵ FRANZ EISENFÜHR/MARTIN WEBER, Rationales Entscheiden, Heidelberg 1993, 169; ROLF BENDER/ARMIN NACK, Tatsachenfeststellung vor Gericht, Bd. 1 Glaubwürdigkeits- und Beweislehre, 2. Aufl. München 1995, Rz. 379.

⁶⁰⁶ AMOS TVERSKY/DANIEL KAHNEMAN, Causal schemas in judgments under uncertainty, in: MARTIN FISHBEIN (Hrsg.), Progress in social psychology, Hillsdale 1980, 49-72; DON LYON/PAUL SLOVIC, Dominance of Accuracy Information and Neglect of Base-Rates in Probability Estimates, Acta Psychologica 1976, 287-298; MAYA BAR-HILLEL, The Base-Rate Fallacy in Probability Judgments, Acta Psychologica 1979, 211-233, 220.

⁶⁰⁷ Für eine andere Lösung nach dem so genannten Signalentdeckungsmodell siehe GERD GIGERENZER, Die Repräsentation von Information und ihre Auswirkung auf statistisches Denken, in: HELL/FIEDLER/GIGERENZER (Hrsg.), FN 11, 99-127, 117 ff. Die Lösung nach dem Signalentdeckungsmodell setzt aber voraus, dass der Zeuge die Verteilung von grünen und blauen Taxis kennt, was kaum je der Fall sein dürfte.

⁶⁰⁸ Möglicherweise zu Unrecht, siehe STEPHEN M. STIGLER, Who discovered Bayes's Theorem?, The American Statistician 1983, 290-296.

II. Der normative Standard: das Bayes-Theorem

365 Das Bayes-Theorem zeigt, wie die anfängliche subjektive Überzeugung, dass eine Hypothese zutrifft (a-priori, Vortest-, Ursprungs- oder Anfangswahrscheinlichkeit) neuen Erkenntnissen anzupassen ist.

366 Wenn H eine Hypothese bezeichnet, I ein Indiz und P „Wahrscheinlichkeit“ (von *probability*), wobei P Wert zwischen 0 und 1 annehmen kann, dann gilt⁶⁰⁹

$$P(H | I) = \frac{P(H) \cdot P(I | H)}{P(I)} \quad (7)$$

367 In Worten ausgedrückt, berechnet sich die bedingte Wahrscheinlichkeit, dass die Hypothese H bei Vorliegen des Indizes I zutrifft, aus der Wahrscheinlichkeit, dass die Hypothese H zutrifft, multipliziert mit der bedingten Wahrscheinlichkeit, dass das Indiz I vorliegt, wenn die Hypothese H zutrifft, geteilt durch die Wahrscheinlichkeit, dass das Indiz I überhaupt vorliegt (unabhängig davon, ob H zutrifft oder nicht). In der Terminologie von ARMIN NACK ist $P(I | H)$ die *Merkmalswahrscheinlichkeit*, während $P(H | I)$ die *Belastungswahrscheinlichkeit* ist.⁶¹⁰

368 Durch Umformung der Gleichung (7) lässt sich das Bayes-Theorem als so genannte „ratio rule“⁶¹¹ darstellen

$$\frac{P(H | I)}{P(I | H)} = \frac{P(H)}{P(I)} \quad (8)$$

369 Mit anderen Worten ist das Verhältnis der bedingten Wahrscheinlichkeiten $P(H | I)$ und $P(I | H)$ gleich dem Verhältnis der einfachen Wahrscheinlichkeiten $P(H)$ und $P(I)$.

370 Für den Fall von zwei sich gegenseitig ausschliessenden Hypothesen, von denen eine zutreffen muss (Beispiel: „A ist der Vater von B“, „A ist nicht der Vater von B“), lässt sich die bedingte Wahrscheinlichkeit, dass die Hypothese bei Vorliegen des Indizes I (Beispiel: „A und B haben die gleiche Blutgruppe, die nur bei zehn Prozent der Bevölkerung vorkommt“), wie folgt errechnen

$$P(H | I) = \frac{P(H) \cdot P(I | H)}{P(H) \cdot P(I | H) + P(\neg H) \cdot P(I | \neg H)} \quad (9)$$

371 In dieser Form wird das Bayes-Theorem oft in der juristischen oder psychologischen Literatur dargestellt.⁶¹² Der Nenner des Bruches auf der rechten Seite der Gleichung, die

⁶⁰⁹ Die Herleitung des Bayes-Theorem aus den Axiomen der Wahrscheinlichkeitstheorie ist in Anhang A dargestellt.

⁶¹⁰ ARMIN NACK, Beweiswürdigung beim Indizienbeweis, Kriminalistik 1995, 466-470, 468.

⁶¹¹ HASTIE/DAWES, FN 399, 121.

Randwahrscheinlichkeit von Indiz I, errechnet sich aus der Wahrscheinlichkeit, dass Indiz I vorliegt, wenn Hypothese H zutrifft und der Wahrscheinlichkeit, dass Indiz I vorliegt, wenn die Hypothese H nicht zutrifft. Im Zusammenhang mit Diagnoseverfahren wird das Vorhandensein von Indiz I, obwohl die Hypothese H nicht zutrifft, als „falsch positiv“ bezeichnet, und die Rate der falsch positiven Befunde ist ein wichtiges Kriterium für die Beweiskraft des Tests (Beispiel: Der Gynäkologe ertastet Knötchen in der Brust einer Patientin, obwohl die Hypothese „Brustkrebs“ nicht vorliegt).

- 372 Leicht verständlicher und praktischer in der Anwendung auf Fragestellungen, wie sie typischerweise Juristen beschäftigen, ist das Bayes-Theorem oft in seiner in „Chancen“ („eins zu zehn“; Englisch: *odds*) ausgedrückten Form:

$$\frac{P(H_1 | I)}{P(H_2 | I)} = \frac{P(H_1)}{P(H_2)} \cdot \frac{P(I | H_1)}{P(I | H_2)} \quad (10)$$

- 373 Der letzte Bruch auf der rechten Seite der Gleichung (10) ist der so genannte „Likelihood-Quotient“⁶¹³, der den abstrakten Beweiswert eines Indizes angibt.⁶¹⁴ In Worten daher

Bedingte Chance = Anfangschance · Likelihood-Quotient

- 374 Angenommen, die Chance, dass der Angeschuldigte sich am Tatort befunden hat, beträgt aufgrund der bisher bekannten Tatsachen 3:4 (dies entspricht einer Wahrscheinlichkeit von 0,42; eine Chance m:n entspricht der Wahrscheinlichkeit $m / (m+n)$, oder 3/7 in diesem Beispiel). Nun trifft ein Laborbericht ein, gemäss dem am Tatort Spuren einer Blutgruppe gefunden wurde, die selten ist und nur in einem Prozent der Bevölkerung vorkommt (Indiz I). Der Angeschuldigte weist die identische Blutgruppe und eine Schnittverletzung auf, während das Opfer eine andere Blutgruppe hat. Wenn H_1 die Hypothese bezeichnet, dass sich der Angeschuldigte am Tatort befunden hat, dann ist die bedingte Wahrscheinlichkeit $P(I | H_1) = 1$. Die bedingte Wahrscheinlichkeit $P(I | H_2)$, d.h. dass die Blutgruppen übereinstimmen, obwohl sich der Angeklagte nicht am Tatort befunden hat (und die Spur daher von jemand anderen verursacht wurde), beträgt 0,01. Es ist demnach 100 Mal wahrscheinlicher, dass die Spuren übereinstimmen, wenn sich der Angeschuldigte am Tatort befunden hat, als wenn er sich nicht am Tatort befunden hat. Der Likelihood-Quotient beträgt daher 100. Multipliziert man die Anfangschance mit dem Likelihood-Quotienten ($3:4 \cdot 100:1 = 300:4$), erhält man eine Wahrscheinlichkeit von 0,986 (300/304), dass sich der Angeschuldigte am Tatort befunden hat.

⁶¹² BENDER/NACK, FN 605, Rz. 401, dort auch Erweiterung auf mehrere Indizien; GÄELLE VILLEJOUBERT/DAVID R. MANDEL, The Inverse Fallacy: An Account of Deviations from Bayes's Theorem and the Additivity Principle, *Memory & Cognition* 2002, 171-178, 171.

⁶¹³ RICHARD JEFFREY, *Subjective Probability: The Real Thing*, Cambridge 2004, 16.

⁶¹⁴ BENDER/NACK, FN 605, Rz. 405; THOMAS D. LYON/JONATHAN J. KOEHLER, The Relevance Ratio: Evaluating the Probative Value of Expert Testimony in Child Sexual Abuse Cases, *Cornell Law Review* 1996, 43-78, 46.

A. Die Lösung des Taxi-Problems nach dem Bayes-Theorem

- 375 Mit dem Bayes-Theorem lässt sich nun die Wahrscheinlichkeit, dass der Unfall von einem grünen Taxi verursacht wurde, wenn der Zeuge aussagt, dass das Taxi grün gewesen sei, leicht berechnen.
- 376 Im Taxi-Problem ist die Anfangswahrscheinlichkeit – ohne zusätzliche Informationen, und unter der Annahme, dass beide Taxi-Gesellschaften gleich viele Unfälle verursachen –, dass ein Taxi der Gesellschaft B den Unfall verursacht hat, 85 % [$P(A) = 0,85$]. Die Anfangswahrscheinlichkeit, dass ein Taxi der Gesellschaft A den Unfall verursacht hat, beträgt 15 %, da es nur zwei Taxigesellschaften gibt. Die bedingte Wahrscheinlichkeit, dass der Zeuge das Taxi als grün identifiziert, wenn das Taxi tatsächlich grün ist (und folglich der Gesellschaft A gehört), beträgt 80 % [$P(\text{„grün“} | A) = 0,8$]; umgekehrt beträgt die Wahrscheinlichkeit, dass der Zeuge das Taxi fälschlicherweise als grün identifiziert, wenn es blau ist und der Gesellschaft B gehört, 20 %.
- 377 Aus diesen Wahrscheinlichkeiten lässt sich die a-posteriori Wahrscheinlichkeit oder Belastungswahrscheinlichkeit, dass das Taxi der Gesellschaft A gehört, wenn der Zeuge ein grünes Taxi gesehen, berechnen:

$$P(A | \text{„grün“}) = \frac{P(A) \cdot P(\text{„grün“} | A)}{P(\text{„grün“})} \quad (11)$$

- 378 Der Nenner dieser Formel bedarf wie bereits erwähnt vorgängiger Berechnung. Er gibt an, wie wahrscheinlich es überhaupt ist, dass der Zeuge „grün“ sagt, und setzt sich daher zusammen aus der Wahrscheinlichkeit, dass der Zeuge „grün“ sagt, wenn das Taxi tatsächlich grün ist (also $P(A) \cdot P(\text{„grün“} | A)$) und der Wahrscheinlichkeit, dass der Zeuge ein Taxi als grün identifiziert, das tatsächlich blau ist (also $P(B) \cdot P(\text{„grün“} | B)$). Folglich lautet die Formel in ihrer extensiven Form

$$P(A | \text{„grün“}) = \frac{P(A) \cdot P(\text{„grün“} | A)}{P(A) \cdot P(\text{„grün“} | A) + P(B) \cdot P(\text{„grün“} | B)} \quad (12)$$

- 379 Setzt man die Wahrscheinlichkeiten gemäss dem Beispiel ein, ergibt sich eine Wahrscheinlichkeit von 41 %, dass ein grünes Taxi den Unfall verursacht hat, wenn der Zeuge ein grünes Taxi gesehen hat.⁶¹⁵ Dieser Wert ist nach der positiven Identifizierung durch den Zeugen zwar erheblich höher als die Anfangswahrscheinlichkeit von 15 %, aber er darf nicht gleichgesetzt werden mit der Wahrscheinlichkeit von 80 %, dass der Zeuge ein grünes Taxi richtig identifiziert. Letzteres, die Gleichsetzung der bedingten Wahrscheinlich-

⁶¹⁵ $(0,15 \cdot 0,8) / ((0,15 \cdot 0,8) + (0,85 \cdot 0,2)) = 0,41$ (gerundet auf zwei Stellen hinter dem Komma). Wer nicht selber rechnen will, kann auf einen der zahlreichen im Internet zugänglichen „Bayes-Rechner“ zurückgreifen, siehe z.B. <http://psych.fullerton.edu/mbirnbaum/bayes/BayesCalc.htm> oder <http://members.aol.com/johnp71/bayes.html> (besucht am 14. März 2004).

keit $P(A \mid \text{„grün“})$ mit $P(\text{„grün“} \mid A)$ ist ein häufiger Fehler, der auf Denken in Assoziationen, oder Repräsentativität, zurückzuführen ist.⁶¹⁶

III. Das deskriptive Modell: Denken in Repräsentativität

- 380 Dass das Bayes-Theorem nicht beschreibt, wie Menschen tatsächlich mit probabilistischen Informationen umgehen, wurde bereits in den 60-er Jahren des letzten Jahrhunderts erkannt. Studien von WARD EDWARDS und Kollegen kamen zum Schluss, dass Menschen ihre subjektive Einschätzung der Wahrscheinlichkeit nach der Kenntnis neuer Indizien zwar in die vom Bayes-Theorem vorgegebene Richtung revidierten, aber zu wenig. Dieses Phänomen wurde daher als „Konservatismus“ bezeichnet, weil die Versuchspersonen an ihren ursprünglichen Überzeugungen festhielten.⁶¹⁷ Die Resultate von EDWARDS waren aber nicht immer eindeutig; so verschwand beispielsweise der Konservatismus, wenn die neue Information nur geringen diagnostischen Wert hatte (der Likelihood-Quotient also nahe bei 1 liegt). In diesem Fall war die Revision der Anfangswahrscheinlichkeit sogar zu extrem.⁶¹⁸ Andere Forscher wiesen darauf hin, dass in der Wirklichkeit viele Indizien redundant seien und daher ein einzelnes Indiz bei der Revision der Anfangswahrscheinlichkeit zu Recht nur gering gewichtet wird. Die Versuchspersonen würden möglicherweise ihre dem tatsächlichen Leben angepassten Entscheidungsalgorithmen auf die künstliche Versuchsanlage übertragen; die Resultate seien daher ein Artefakt der Versuchsanlage. Auch sei die Diagnostizität der Indizien unrealistisch hoch; die Versuchspersonen würden automatisch von realistischeren Werten ausgehen.⁶¹⁹ Diese Kritik koinzidierte mit der Publikation von „*On the Psychology of Prediction*“ von KAHNEMAN und TVERSKY, in der sie nachwiesen, dass ihre Versuchspersonen Basisraten vernachlässigen oder gar ignorieren.⁶²⁰ Die Entdeckung der Vernachlässigung der Anfangswahrscheinlichkeit (*base rate neglect*) – der Antithese zum Konservatismus – bedeutete zusammen mit der Kritik an der künstlichen Versuchsanordnung das Ende der Forschung zum Konservatismus.⁶²¹

⁶¹⁶ HASTIE/DAWES, FN 399, 127; ROBYN M. DAWES, Representative Thinking in Clinical Judgment, *Clinical Psychology Review* 1986, 425-441, 430; EISENFÜHR/WEBER, FN 605, 171 f.; ARMIN NACK, FN 610, 469; VILLEJOUBERT/MANDEL, FN 605, 17.

⁶¹⁷ WARD EDWARDS, Conservatism in human information processing, in B. KLEINMUNTZ (Hrsg.), *Formal Representation of Human Judgement*, New York 1968; gekürzter Nachdruck in: KAHNEMAN/SLOVIC/TVERSKY (Hrsg.), 359-369.

⁶¹⁸ LAWRENCE D. PHILLIPS/WARD EDWARDS, Conservatism in simple probability inference tasks, *Journal of Experimental Psychology* 1966, 346-357.

⁶¹⁹ PETER AYTON/GEORGE WRIGHT, Subjective Probability: What Should We Believe?, in: GEORGE WRIGHT/PETER AYTON (Hrsg.), *Subjective Probability*, New York 1994, 163-183, 168 f.

⁶²⁰ DANIEL KAHNEMAN/AMOS TVERSKY, On the Psychology of Prediction, *Psychological Review* 1973, 237-251; Nachdruck in: KAHNEMAN/SLOVIC/TVERSKY (Hrsg.), 48-68.

⁶²¹ AYTON/WRIGHT, FN 619, 170. Was nicht heisst, dass das Phänomen nicht existiert. Dort, wo die neue Information sehr hohe Diagnostizität hat, scheinen Menschen ihre Anfangswahrscheinlichkeit zu wenig stark zu revidieren; so WILLIAM C. THOMPSON/EDWARD L. SCHUMANN, Interpretation of Statistical Evidence in Criminal Trials: The Prosecutor's Fallacy and the Defense Attorney's Fallacy, *Law and Human Behavior* 1987, 167-187, 175; DAVID L. FAIGMAN/A.J. BAGLIONI, Bayes' Theorem in the Trial Process: Instructing Jurors on the Value of Statistical Evidence, *Law and Human Behavior* 1988, 1-17, 13.

- 381 KAHNEMAN und TVERSKY postulieren, dass Menschen sich einer Repräsentativitätsheuristik (*representative heuristic*) bedienen, wenn sie die Wahrscheinlichkeit des Vorliegens einer Tatsache einschätzen. Gemäss der Repräsentativitätsheuristik beurteilt eine Person die Wahrscheinlichkeit eines ungewissen Ereignisses nach dem Ausmass, in dem es in wesentlichen Eigenschaften seiner Grundgesamtheit ähnlich ist.⁶²² Mit anderen Worten postuliert die Repräsentativitätsheuristik, dass der Mensch auch dort assoziativ denkt, wo er komparativ denken sollte. Beim assoziativen Denken werden wahrgenommene Stimuli mit den Eigenschaften von im Gedächtnis gespeicherten Schemata verglichen. Ein Schema ist eine komplexe Wissenseinheit, die die typischen Eigenschaften von Mitgliedern einer allgemeinen Kategorie enthält.⁶²³ Nachdem das Schema aktiviert ist, können daraus weitere, noch nicht beobachtete Eigenschaften des betreffenden Objekts abgeleitet werden. Dies ist häufig nützlich; wird beispielsweise ein Tier mit einer gewissen Grösse, Farbe und Form wahrgenommen, kann es dem Schema „Stinktief“ zugeordnet werden. Ohne wahrnehmen zu müssen, dass das Tier tatsächlich stinkt, kann die Flucht ergriffen werden.⁶²⁴
- 382 Unter gewissen Umständen kann das Denken in Assoziationen zu Fehlern führen, wie folgendes klassisches Experiment zeigt. KAHNEMAN und TVERSKY legten ihren Versuchspersonen fünf Beschreibungen von Menschen vor, die zufällig aus einer Gruppe von 100 Personen gezogen worden waren. Der einen Hälfte der Versuchspersonen wurde gesagt, dass die Gruppe aus 70 Ingenieuren und 30 Anwälten, der anderen Hälfte, dass sie aus 30 Ingenieuren und 70 Anwälten bestehe. Ansonsten waren die Instruktionen identisch.
- 383 Eine solche Beschreibung lautete beispielsweise:⁶²⁵
- Jack ist ein 45-jähriger Mann. Er ist verheiratet und hat vier Kinder. Er ist meist konservativ, sorgfältig und ehrgeizig. Er zeigt kein Interesse an politischen und sozialen Themen und verbringt seine Freizeit mit seinen zahlreichen Hobbies, die Heimwerken, Segeln und das Lösen mathematischer Rätsel umfassen.
- Wie gross ist die Wahrscheinlichkeit, dass Jack einer der 30 Ingenieure aus der Gruppe der 100 Personen ist?
- 384 Die Antwort sollte, unabhängig davon, wie diagnostisch die Information ist,⁶²⁶ dadurch beeinflusst werden, wie die gross die Anfangswahrscheinlichkeit ist, dass Jack ein Ingenieur ist. Angenommen, eine Versuchsperson ist aufgrund der Beschreibung der Meinung, dass Männer mit den beschriebenen Eigenschaften unter Ingenieuren fünf Mal häufiger anzutreffen sind als unter Anwälten. In diesem Fall ist die a posteriori Wahrscheinlichkeit,

⁶²² DANIEL KAHNEMAN/AMOS TVERSKY, Subjective probability: A Judgment of Representativeness, *Cognitive Psychology* 1972, 430-454, 431; Nachdruck in: KAHNEMAN/SLOVIC/TVERSKY (Hrsg.), 32-47, 33. Später haben KAHNEMAN und TVERSKY die Definition dessen, was „repräsentativ“ bedeutet, erheblich ausgeweitet, AMOS TVERSKY/DANIEL KAHNEMAN, Judgments of and by Representativeness, in: KAHNEMAN/SLOVIC/TVERSKY (Hrsg.), 84-100, 87; nicht ganz zu Unrecht wurde ihnen vorgeworfen, dass die Definition sehr vage und offen wird, MARGRIT OSWALD, Urteile über den Repräsentativitätsheurismus, *Archiv für Psychologie* 1986, 113-125, 114 ff.

⁶²³ EISENFÜHR/WEBER, FN 605, 171.

⁶²⁴ Beispiel aus HASTIE/DAWES, FN 399, 117.

⁶²⁵ KAHNEMAN/TVERSKY, FN 620, 54.

⁶²⁶ Die Kritik von OSWALD, FN 622, 122, am „Tom W.“ Experiment (KAHNEMAN/TVERSKY, FN 620, 238) ist zwar berechtigt, vermag aber nicht die Resultate des „Ingenieur/Anwalt“ Experiments zu erklären.

dass Jack ein Ingenieur ist, 69 %, ⁶²⁷ wenn sich 30 Ingenieure in der Gruppe befinden, wenn sich jedoch 70 Ingenieure in der Gruppe der 100 befinden, beträgt sie 92 %. ⁶²⁸

- 385 KAHNEMAN und TVERSKY stellten jedoch keine signifikanten Unterschiede zwischen den beiden Gruppen fest. Ihre Versuchspersonen hatten die Anfangswahrscheinlichkeit schlicht ignoriert; beachtet wurde sie nur dann, wenn überhaupt keine individualisierende Information vorhanden war („Sie haben keinerlei Information über die Person, die zufällig aus der Stichprobe gezogen wurde. Wie gross ist die Wahrscheinlichkeit, dass die Person einer der 30 Ingenieure in der Gruppe der 100 Personen ist?). ⁶²⁹ Andere Autoren replizierten und erweiterten die Experimente von KAHNEMAN und TVERSKY und bestätigten die Resultate weitgehend. ⁶³⁰ Der *base rate neglect* war geboren und führte zu einer Flut von Publikationen, die bis heute nicht nachlässt (zur Kritik am *base rate neglect* siehe hinten, S. 139 ff.).
- 386 Die Repräsentativität einer Eigenschaft einer Instanz einer Kategorie für ihre Zugehörigkeit zu derselben kann durch die bedingte Wahrscheinlichkeit $P(E | S)$ ausgedrückt werden; ⁶³¹ wobei „E“ die Eigenschaft und „S“ das Schema bezeichnet. Je repräsentativer, oder typischer, die Eigenschaften für die betreffende Kategorie sind, desto höher ist die bedingte Wahrscheinlichkeit $P(E | S)$; sie erreicht 1, wenn alle Instanzen der Kategorie die entsprechende Eigenschaft aufweisen. In dem „Jack der Ingenieur“ Experiment haben möglicherweise viele Versuchspersonen assoziativ gedacht: „Das ist ein typischer Ingenieur. Mindestens 50 % aller Ingenieure haben diese Eigenschaften“. Die bedingte Wahrscheinlichkeit $P(E | S)$ beträgt daher 0,5; d.h. die Wahrscheinlichkeit, dass die Eigenschaften gegeben sind, wenn eine Instanz des Schemas „Ingenieur“ vorliegt, ist 50 %. Falsch ist jedoch der nächste Schritt: „Weil 50 % aller Ingenieure diese Eigenschaften aufweisen, ist die Wahrscheinlichkeit, dass Jack ein Ingenieur ist, 50 %“. $P(E | S)$ darf, siehe *ratio rule* (Gleichung 8, S. 125), nur dann mit $P(S | E)$ gleichgesetzt werden, wenn die Wahrscheinlichkeit von S und E identisch sind. Repräsentatives Ähnlichkeits-Denken unterscheidet jedoch nicht zwischen $P(E | S)$ und $P(S | E)$ und führt eine Symmetrie ein, die in Wirklichkeit (meist) nicht existiert. ⁶³²
- 387 Die Umkehrung bedingter Wahrscheinlichkeiten (*inverse fallacy* ⁶³³) ist sehr häufig und hat VILLEJOUBERT/MANDEL dazu verleitet, anzunehmen, dass sie die Ursache des empirisch festgestellten *base rate neglects* ist. ⁶³⁴ Während bei der Vernachlässigung der Anfangswahrscheinlichkeit nur die Basisrate ignoriert wird, wird bei der Verwechslung der bedingten Wahrscheinlichkeiten auch $P(I | \neg H)$ vernachlässigt (im Beispiel von „Jack“ also

⁶²⁷ Nach dem Bayes-Theorem in seiner Chancen-Form: $3:7 \cdot 5 = 15:7$ ($P = 0,68$).

⁶²⁸ $7:3 \times 5 = 35:3$ ($P=0,92$).

⁶²⁹ KAHNEMAN/TVERSKY, FN 620, 241.

⁶³⁰ MAYA BAR-HILLEL, Studies of Representativeness, in: KAHNEMAN/SLOVIC/TVERSKY (Hrsg.), 69-83; MAYA BAR-HILLEL/EFRAT NETER, How Alike Is It? Versus How Likely Is It?: A Disjunction Fallacy in Probability Judgements, Journal of Personality and Social Psychology 1986, 1119 -1131; Nachdruck in: GILOVICH/GRIFFIN/KAHNEMAN (Hrsg.), 82-97.

⁶³¹ HASTIE/DAWES, FN 399, 121.

⁶³² HASTIE/DAWES, FN 399, 122.

⁶³³ JONATHAN J. KOEHLER, The Base Rate Fallacy Reconsidered: Descriptive, Normative and Methodological Challenges, Behavioral and Brain Sciences 1996, 1-53, 9.

⁶³⁴ VILLEJOUBERT/MANDEL, FN 612, 171 ff.

die Wahrscheinlichkeit, dass eine Person mit den beschriebenen Eigenschaften auch unter Anwälten vorkommt).

- 388 VILLEJOUBERT/MANDEL machen sich bei ihrem Experiment die Tatsache zunutze, dass es auf die Basisraten nicht ankommt, wenn die Anfangswahrscheinlichkeit 0,5 ist. Sie legten ihren Versuchspersonen insgesamt zwölf Fragen im folgenden Format vor: „Auf einem Planeten gibt es eine Million Grooms und eine Million Fizes. 98 % der Grooms und 58 % der Fizes spielen Harmonika. Sie treffen eine der Kreaturen, die Ihnen durch einen Übersetzer mitteilt, dass sie Harmonika spielt. Wie gross ist die Wahrscheinlichkeit, dass es sich um einen Groom handelt?“. VILLEJOUBERT/MANDEL stellten fest, dass die Antworten häufig näher bei $P(I | H)$ (im obigen Beispiel 0,98) als bei der gemäss Bayes-Theorem korrekten Antwort lagen (im obigen Beispiel 0,63). Rund die Hälfte ihrer Versuchspersonen setzten in mehr als 80 % der Fragen a posteriori Wahrscheinlichkeit und bedingte Wahrscheinlichkeit $P(I | H)$ gleich.⁶³⁵ Für VILLEJOUBERT/MANDEL ist dies Beleg dafür, dass nicht die Vernachlässigung der Basisraten, sondern die Verwechslung der bedingten Wahrscheinlichkeiten die Ursache für die beobachteten Abweichungen vom Bayes-Theorem ist.
- 389 Die Studie von VILLEJOUBERT/MANDEL ist ein starkes Indiz für assoziatives Denken, wenn komparatives Denken angezeigt wäre. KAHNEMAN/TVERSKY behaupten allerdings nicht, dass die Vernachlässigung der Basisraten die Ursache der beobachteten Abweichungen von der gemäss Bayes-Theorem richtigen Antwort ist; Ursache ist gemäss ihnen das repräsentative Denken. Die Vernachlässigung der Basisraten ist nur *eine* der zahlreichen Folgen dieser Heuristik.⁶³⁶ Allerdings hat sie in den letzten zwanzig Jahren die Diskussion dominiert, weshalb häufig nur noch von *base rate neglect* gesprochen wird, was VILLEJOUBERT/MANDEL möglicherweise zur Annahme verleitet hat, die Vernachlässigung der Basisrate werde als Ursache der beschriebenen Fehler gesehen.

A. Typische Fehler, die aus dem repräsentativen Denken resultieren

1. Ausländer sind kriminell, Marihuana ist eine Einstiegsdroge und Videospiele machen gewalttätig

- 390 Der Anteil ausländischer Staatsangehöriger in den Strafvollzugsanstalten des Kantons Zürich (ohne Ausschaffungshaft) betrug 2003 76,5 %.⁶³⁷ Praktisch alle jugendlichen Amokläufer – der Attentäter von Erfurt, die Mörder von Columbine, Colorado – spielten intensiv gewalttätige Videospiele. Und fast alle Konsumenten von harten Drogen konsumieren auch, oder konsumierten früher, Marihuana. Dies hat Leute dazu verleitet, darauf

⁶³⁵ VILLEJOUBERT/MANDEL, FN 612, 176.

⁶³⁶ Weitere Folgen sind beschrieben in KAHNEMAN/FREDERICK, FN 245.

⁶³⁷ Justizvollzug des Kantons Zürich, Jahresheft 2004, 64 (erhältlich unter www.poeschwies.ch/content/pdf/suche/index.php?id_pdf=395; besucht am 10. Januar 2005).

zu schliessen, dass Ausländer generell kriminell sind,⁶³⁸ Videospiele zu Gewalttaten führen⁶³⁹ und Marihuana eine Einstiegsdroge ist.⁶⁴⁰

- 391 Diese Schlüsse sind falsch, weil sie aus der bedingten Wahrscheinlichkeit $P(I | H)$ auf die bedingte Wahrscheinlichkeit $P(H | I)$ schliessen. Um das Verhältnis der bedingten Wahrscheinlichkeiten bestimmen zu können, muss man aber auch das Verhältnis Wahrscheinlichkeiten $P(H)$ und $P(I)$ kennen. Um beim Beispiel der Videospiele zu bleiben: $P(H)$, d.h. Amokläufer, ist extrem selten. Selbst in den USA, wo solche Amokläufe häufiger sind als in anderen Ländern, kommen sie kaum mehr als ein Mal jährlich vor. Der Anteil Jugendlicher, zumal männlicher, die Videospiele spielen, ist aber sehr hoch.⁶⁴¹ Gewalttätige Videospiele gehören dabei zu den populärsten.⁶⁴² Da das Verhältnis $P(\text{„läuft Amok“})$ zu $P(\text{„spielt Videogames“})$, sehr gross ist, muss auch das Verhältnis $P(\text{„läuft Amok“} | \text{„spielt Videogames“})$ zu $P(\text{„läuft nicht Amok“} | \text{„spielt Videogames“})$ sehr gross sein. Mit anderen Worten wird nur ein ganz, ganz kleiner Anteil der Videospieleler zu Gewalttätern, so wie nur ein kleiner Anteil der Haschisch-Konsumenten zu harten Drogen greift.⁶⁴³

2. Fallacia Consequentis

- 392 Mit den obigen Fehlschlüssen verwandt ist die bereits Aristoteles bekannte *fallacia consequentis*, der bejahende Schluss von der Folge auf den Grund.⁶⁴⁴ Die *fallacia consequentis* hat folgende Form:

Wenn p, dann q.	(Wer gewalttätig ist, spielt brutale Videospiele.)
q.	(X spielt brutale Videospiele.)
Ergo, p.	(X ist gewalttätig.)

- 393 Selbst wenn die Aussage „wenn p, dann q“ wahr ist – die bedingte Wahrscheinlichkeit $P(q | p)$ also 1 ist – ist der obige Schluss nicht richtig.⁶⁴⁵ Zutreffend ist der Schluss nur dann, wenn die Prämisse lautet „Ausschliesslich wenn p, dann q“; d.h. das Indiz kommt nur bei der Haupttatsache vor, und sonst gar nicht. In Bayes'scher Terminologie ist der Likeli-

⁶³⁸ Tatsächlich sind die Unterschiede zwischen Ausländern und Schweizern gering, wenn man die unterschiedliche Alters- und Geschlechtsstruktur berücksichtigt. Betrachtet man ausschliesslich die Wohnbevölkerung (lässt also Asylanten und Touristen oder Durchreisende weg), so fehlen die Unterschiede sogar fast ganz; Bundesamt für Statistik, Pressemitteilung vom 10. Mai 1996.

⁶³⁹ „Software für Massaker“, Frankfurter Allgemeine Sonntagszeitung vom 28. April 2002.

⁶⁴⁰ Dieses und weitere Beispiele bei HASTIE/DAWES, FN 399, 122.

⁶⁴¹ DOUGLAS A. GENTILE/DAVID A. WALSH, A Normative Study of Family Media Habits, Applied Developmental Psychology 2002, 157-178.

⁶⁴² CRAIG A. ANDERSON/BRAD J. BUSHMAN, Effects of Violent Video Games on Aggressive Behavior, Aggressive Cognition, Aggressive Affect, Physiological Arousal and Prosocial Behavior: A Meta-Analytic Review of the Scientific Literature, Psychological Science 2001, 353-359.

⁶⁴³ BVerfGE 90, 145, 181 (in Deutschland rund 2,5%).

⁶⁴⁴ EGON SCHNEIDER, Logik für Juristen, 5. Aufl. München 1999, 259.

⁶⁴⁵ Zu den häufigsten Fehlern beim deduktiven Schliessen siehe VOLKER GADENNE, Deduktives Denken und Rationalität, in: HELL/FIEDLER/GIGERENZER (Hrsg.), FN 11, 161-188, 163 ff.; weiteres Beispiel bei FRANCO TARONI/PATRICE MANGIN/WALTER BÄR, Die Interpretation des Beweiswertes von DNA-Untersuchungen in sachverständigen Gutachten, Schweizerische Zeitschrift für Strafrecht 1999, 439-445, 440.

hood-Quotient für dieses Indiz unendlich, sein abstrakter Beweiswert ist unermesslich.⁶⁴⁶ Ergo führt es selbst dann zur Bejahung von p, wenn die Anfangswahrscheinlichkeit von p sehr gering ist.

3. Was typisch ist, ist beweiskräftig

- 394 Beweisanzeichen, die typisch sind für eine bestimmte Tat, müssen nicht notwendigerweise diagnostisch sein. Kinder, die über sexuelle Übergriffe berichten, zögern oft und widerrufen bereits gemachte Aussagen.⁶⁴⁷ Die Zurücknahme einer Anschuldigung ist daher nicht untypisch für ein missbrauchtes Kind und Teil von SUMMITs „*Child Sexual Abuse Accommodation Syndrome*“.⁶⁴⁸
- 395 Nur, wie gross ist die Wahrscheinlichkeit, dass ein Kind, das *nicht* missbraucht wurde und beispielsweise durch suggestive Fragen dazu gebracht wurde, unrichtige Anschuldigungen zu erheben,⁶⁴⁹ diese Aussagen später widerruft? Der gesunde Menschenverstand sagt einem, dass die Rate der widerrufenen Anschuldigungen in diesem Fall höher sein muss. Der Widerruf einer Anschuldigung muss daher ein entlastendes Indiz sein. Wie entlastend, können nur empirische Untersuchungen zeigen, die eine Kontrollgruppe verwenden und den Vergleich der Wahrscheinlichkeit eines Widerrufs bei tatsächlichem Missbrauch und suggeriertem Missbrauch erlauben.
- 396 Auch andere Anzeichen, die oft als „typisch“ für Kindsmissbrauch angesehen werden, taugen nicht als Indizien für Missbrauch. So ist eine Rötung der Genitalien bei missbrauchten Kindern zwar häufig, aber bei nicht missbrauchten ebenso.⁶⁵⁰ Auch psychosomatische Symptome wie Bettnässen, Kopfschmerzen oder Verstopfung treten bei missbrauchten Kindern nicht häufiger auf als bei nicht missbrauchten Kindern in vergleichbarer Situation.⁶⁵¹ Von der Typizität der Symptome auf die Ursache zu schliessen, ohne die Häufigkeit der Symptome bei Nicht-Vorliegen der Ursache zu kennen, ist ein typischer Fehler assoziativen Denkens.
- 397 Das Problem ist, dass häufig Experten befragt wurden, die ausschliesslich mit missbrauchten Kindern zu tun hatten. Ihre Beobachtungen aber können nie dazu dienen, zwischen den Hypothesen Missbrauch und Nicht-Missbrauch zu differenzieren, auch wenn das zu beur-

⁶⁴⁶ BENDER/NACK, FN 605, Rz. 439.

⁶⁴⁷ TEENA SORENSEN/BARBARA SNOW, How Children Tell: The Process of Disclosure in Child Sexual Abuse, Child Welfare 1991, 3-15, zitiert in DEBRA POOLE/D. STEPHEN LINDSAY, Assessing the accuracy of young children's reports: Lessons from the Investigation of Child Sexual Abuse, Applied & Preventive Psychology 1998, 1-26, 9.

⁶⁴⁸ POOLE/LINDSAY, FN 647, 7.

⁶⁴⁹ Diese Wahrscheinlichkeit ist, bei unprofessionell geführten Interviews, nicht unbeträchtlich, siehe die bei POOLE/LINDSAY, FN 648, 4, erwähnten Studien.

⁶⁵⁰ Siehe die bei LYON/KOEHLER, FN 614, Fn. 36 zitierten Studien.

⁶⁵¹ POOLE/LINDSAY, FN 647, 9.

teilende Kind alle Stereotypen des typischen Missbrauchs erfüllt. ROBYN DAWES formuliert dies so.⁶⁵²

Unglücklicherweise sagen viele Sachverständige auf der Basis ihrer „Erfahrung“ zu den Merkmalen missbrauchter Kinder aus, aber sie geben zu, wenig oder keine Erfahrungen mit Kindern zu haben, die nicht missbraucht wurden, aber einmal ausgesagt hatten, missbraucht worden zu sein. Daher ist solche Erfahrung – nicht bloss „genau gesagt“, sondern fundamental – *irrelevant* für die Beurteilung des Missbrauchs. Ironischerweise versuchen solche Pseudo-Experten häufig, ihre angeblichen Sachkenntnisse durch die Behauptung zu belegen, dass fast alle Kinder, die sie behandeln, tatsächlich missbraucht worden seien. Wird diese Aussage ernst genommen, sollte sie automatisch dazu führen, ihnen jeden Sachverstand bezüglich der für das Gericht zentralen Frage, nämlich ob das Kind missbraucht wurde oder nicht, abzusprechen.

398 Als Indizien für oder gegen den Missbrauch taugen nur Symptome, deren Häufigkeit bei nicht missbrauchten Kindern in vergleichbaren Umständen man ebenfalls kennt. Dies wurde inzwischen erkannt, und entsprechende Forschungsanstrengungen werden unternommen.⁶⁵³

399 Der Jurist kann daraus lernen, dass sich der Beweiswert eines Indizes immer erst aus Beantwortung der drei folgenden Fragen ergibt:

- a) Wie häufig kommt das Indiz bei der Haupttatsache vor?
- b) Wie häufig kommt das Indiz (auch) bei der Nicht-Haupttatsache vor?
- c) Wo kommt das Indiz häufiger vor, bei Gruppe 1 oder bei Gruppe 2?⁶⁵⁴

400 Lässt sich die Frage 2 nicht beantworten, weil keine Daten vorliegen, so ist das Indiz wissenschaftlich wertlos. Aus den Antworten zu den drei Fragen errechnet sich der abstrakte Beweiswert des Indizes – der Likelihood-Quotient gemäss Bayes-Theorem. BENDER/NACK umschreiben den Beweiswert eines Indizes mit gegebenem Likelihood-Quotienten wie folgt:⁶⁵⁵

1 : 3	gering belastend
1 : 4-9	belastend
1 : 10-25	stark belastend
1 : 50-100	sehr stark belastend
1 : 1'000 und höher	ausserordentlich stark belastend

⁶⁵² Zitiert in POOLE/LINDSAY, FN 647, 21; siehe auch ROBYN DAWES, Probabilistic Thinking, in NEIL J. SMELSER/PAUL B. BALTES (Hrsg.), International Encyclopedia of the Social and Behavioral Sciences, London 2001, 12082-12089, 12086.

⁶⁵³ POOLE/LINDSAY, FN 647, 9.

⁶⁵⁴ NACK, FN 610, 467.

⁶⁵⁵ BENDER/NACK, FN 605, 1. Aufl. München 1981, Rz. 429.

4. Der Trugschluss des Anklägers

- 401 Ein Laden wird überfallen. Der Täter trägt eine Skimütze mit Sehschlitzen. Bei einer Nahbereichsfahndung wird ein Mann aufgegriffen, der ungefähr die Grösse des Räubers hat und ähnliche Kleidung trägt. Weder Beute, Waffe noch Skimütze werden bei ihm gefunden. Eine ähnliche Skimütze, wie sie der Täter trug, findet sich jedoch in einem Abfalleimer unweit des aufgegriffenen Mannes. Nach der Schilderung dieses Sachverhalts – ohne weitere Indizien – wurden die Versuchspersonen in einem Experiment von THOMPSON/SCHUMANN gebeten, ihre subjektive Überzeugung anzugeben, dass es sich bei dem aufgegriffenen Mann um den Täter handle.⁶⁵⁶
- 402 Anschliessend erhielten die Versuchspersonen folgende zusätzliche Informationen: In der Skimütze werden Haare gefunden. Mikroskopische Untersuchungen zeigen, dass sie sich nicht von den Haaren des Verdächtigen unterscheiden lassen. Ein forensischer Wissenschaftler gibt folgende Aussage zu Protokoll:
- wenn der Verdächtige unschuldig wäre, betrüge die Wahrscheinlichkeit, dass die Haarproben nicht unterschieden werden können, nur 2 %
- 403 Nun wurden die Versuchspersonen erneut gebeten, ihre subjektive Überzeugung anzugeben, dass der Verdächtige der Räuber ist. 22 % der Versuchspersonen gaben eine Wahrscheinlichkeit von 98 % an. Sie begingen den Fehler, $P(H | I)$ gleichzusetzen mit $1 - P(I | H)$.⁶⁵⁷ Dieser Fehler wird gewöhnlich als „Trugschluss des Anklägers“ bezeichnet.⁶⁵⁸
- 404 Präsentiert der Sachverständige jedoch seine Erkenntnisse in der Form von Häufigkeiten
- nur 2 % aller Menschen haben Haare, die sich mikroskopisch nicht von den Haaren des Täters unterscheiden lassen. In einer Stadt mit einer Million Einwohner sind dies rund 20'000 Personen
- dann begehen nur 4 % der Befragten den Fehler, $P(H | I)$ gleichzusetzen mit $1 - P(I | H)$.⁶⁵⁹
- 405 Trugschlüsse des Anklägers sind kein Artefakt von gesuchten Experimenten. TARONI/MANGIN/BÄR weisen darauf hin, dass sich falsche Argumente wie das folgende häufig in Expertengutachten finden:⁶⁶⁰
- Es lassen sich keine Unterschiede zwischen den Merkmalen in der untersuchten Spur und denjenigen bestimmt beim Verdächtigen auffinden. Eine solche Befundkonstellation lässt sich bei ca. 0,001 % der Bevölkerung nachweisen. Die Wahrscheinlichkeit, dass die Spur vom Tatverdächtigen stammt, ist grösser als 99,999 %.
- 406 In einem Grundsatzurteil zur DNA-Analyse befasste sich der deutsche Bundesgerichtshof mit den folgenden Aussagen eines Sachverständigen: Die Kombination der Merkmale aller drei DNA-Polymorphismen komme nur bei 0,014 Prozent der Bevölkerung, d.h. bei einer

⁶⁵⁶ THOMPSON/SCHUMANN, FN 621, 173.

⁶⁵⁷ THOMPSON/SCHUMANN, FN 621, 174.

⁶⁵⁸ DAVID J. BALDING/PETER DONNELLY, The Prosecutor's Fallacy and DNA Evidence, Criminal Law Review 1994, 711-721; THOMPSON/SCHUMANN, FN 621, 171; IAN STEWART, Der Trugschluss des Ermittlers, Spektrum der Wissenschaft Juli 1997, 8.

⁶⁵⁹ THOMPSON/SCHUMANN, FN 621, 174.

⁶⁶⁰ TARONI/MANGIN/BÄR, FN 645, 440.

von 6'937 Personen, vor. Deshalb könne mit einer „Wahrscheinlichkeit von 99,986 Prozent (100 Prozent abzüglich 0,014 Prozent)“ festgestellt werden, dass der Angeklagte der Verursacher des Spermas aus der Scheide der Geschädigten sei. Zu dieser Aussage konnte der Gutachter aber nur gelangen, weil er – wie das Gutachter oft tun,⁶⁶¹ aber nicht tun sollten⁶⁶² – von einer „neutralen“ Anfangswahrscheinlichkeit von 0,5 ausging, dass das Sperma vom Angeklagten stammt. Die Feststellung der Belastungswahrscheinlichkeit aus der Merkmalswahrscheinlichkeit setzt aber die Berücksichtigung der Anfangswahrscheinlichkeit voraus. Dass diese 0,5 beträgt, durfte der Sachverständige annehmen, nicht aber das Gericht. Der BGH wies ausdrücklich darauf hin, dass die vom Sachverständigen genannte Merkmalswahrscheinlichkeit von 0,014 bei einer Stadt der Grösse von Hannover mit ungefähr 250'000 Einwohnern immerhin bedeutet, dass 35 männliche Personen aus Hannover als Täter nicht ausgeschlossen werden können.⁶⁶³

- 407 Auch das Bezirksgericht Zürich, so lässt sich aus der Begründung eines Urteils vom 7. Dezember 1993 schliessen, ist nicht gefeit gegen den Trugschluss des Anklägers (der Trugschluss hat in diesem Fall mit an Sicherheit grenzender Wahrscheinlichkeit nicht zu einem Fehlurteil geführt). Das Gericht stellt fest, dass gemäss Expertise der am Tatort gefundene Speichel mit einer Wahrscheinlichkeit von 1 : 150 Millionen vom Angeklagten stamme.⁶⁶⁴ Diese Belastungswahrscheinlichkeit konnte der Experte aber nur angeben, weil er von der neutralen Anfangswahrscheinlichkeit von 50 % ausgegangen ist, dass der Angeklagte der Verursacher der Spuren ist.⁶⁶⁵ Die Annahme einer Anfangswahrscheinlichkeit von 50 % für die Schuld des Angeklagten ist aber einigermassen willkürlich und kann sich zu Ungunsten des Angeklagten auswirken.⁶⁶⁶ Der Gutachter kann die tatsächliche Anfangswahrscheinlichkeit gar nicht kennen,⁶⁶⁷ oft kann sie auch nur nach richterlichem Ermessen geschätzt werden.
- 408 Das Gericht selber ging offensichtlich davon aus, dass die tatsächliche Anfangswahrscheinlichkeit erheblich weniger als 50 % betrug, und schrieb: „Bereits eine erheblich kleinere Wahrscheinlichkeit [als die von 1 : 150 Millionen] würde genügen, um im Angeklagten den Spurenger zu sehen, wenn man bedenkt, dass die Schweiz knapp 7 Millionen, Italien rund 58 Millionen Einwohner zählt“. ⁶⁶⁸ Damit scheint das Gericht andeuten zu wollen, dass die Übereinstimmung bei einer Merkmalswahrscheinlichkeit von 1 : 150 Millionen den Angeklagten selbst dann mit hoher Wahrscheinlichkeit zum Täter macht, wenn jeder der 65 Millionen Einwohner der Schweiz und Italiens als Täter in Frage kommt.

⁶⁶¹ WALTER BÄR, Zum Beweiswert der DNA-Analyse, in: ANDREAS DONATSCH/NIKLAUS SCHMID (Hrsg.), Festschrift für Jörg Rehberg, Zürich 1996, 41-51, 46.

⁶⁶² FRANCO TARONI/COLIN AITKEN, Probabilités et preuve par l'ADN dans les affaires civiles et criminelles: Questions de la cour et réponses fallacieuses des experts, Schweizerische Zeitschrift für Strafrecht 1998, 291-313, 299 f.

⁶⁶³ BGHSt 38, 320.

⁶⁶⁴ ZR 94 Nr. 7, 25 (1995).

⁶⁶⁵ Zur Rechtfertigung dieser Annahme des Gutachters siehe BÄR, FN 661, 46.

⁶⁶⁶ TARONI/AITKEN, FN 662, 299 f.; BGHSt 38, 320.

⁶⁶⁷ BÄR, FN 661, 46; NACK, FN 610, 469.

⁶⁶⁸ ZR 94 (1995) Nr. 7, 25 f.

- 409 Dies ist falsch. Geht man von einer Anfangswahrscheinlichkeit von 1 : 65 Millionen aus, dass der Angeklagte der Spurenverursacher ist, so resultiert aus der Merkmalswahrscheinlichkeit von 1 : 150 Millionen „nur“ eine Wahrscheinlichkeit von 70 %, dass der Angeklagte der Spurenverursacher ist. Nimmt man weiter an, dass es beim Vergleich von zwei DNA-Proben mit einer geringen Wahrscheinlichkeit von 1 : 10'000 zu einem falsch positiven Resultat⁶⁶⁹ kommt – eine durchaus realistische Annahme⁶⁷⁰ – so reduziert sich die Belastungswahrscheinlichkeit dramatisch auf weniger als ein Promille (näheres hinten, S. 170 ff.)⁶⁷¹.
- 410 Das Urteil des Bezirksgerichts Zürich ist trotzdem kein Fehlurteil, weil die Anfangswahrscheinlichkeit, dass der Angeklagte der Spurenverursacher war, im konkreten Fall erheblich höher als 1 : 65 Millionen war. Der dem Angeklagten angelastete Einbruchdiebstahl wurde nach dem gleichen Modus Operandi wie frühere Einbrüche des Angeklagten ausgeführt, der Mittäter war ein alter Bekannter des Angeklagten und der Angeklagte konnte für seinen angeblichen Aufenthalt in Italien keinen einzigen Zeugen benennen.⁶⁷²
- 411 Fälle, in denen Gerichte oder Gerichtsgutachter erwiesenermassen dem Trugschluss des Anklägers unterlagen – und meist von der Berufungsinstanz korrigiert wurden – sind *Regina v. Andrew Deen*, Court of Appeals (Criminal Division) vom 10. Januar 1994; *People v. Collins*, 66 Cal. Rptr. 497 (1968), *U.S. v. Massey*, 594 F. 2d 676 (8th Cir. 1979); *State of Arizona v. Michael Steven Gallegos*, 870 P.2d 1097 (1994) und *State of Arizona v. Robert Wayne Johnson*, 905 P.2d 1002 (1995).

5. Umgekehrter Fehler: was untypisch ist, ist unwahrscheinlich

- 412 ALAN M. DERSHOWITZ, Harvard Professor und Berater des Verteidigungsteams im O.J. Simpson Prozess, sagte im März 1995 am amerikanischen Fernsehen, dass nur etwa ein Zehntel eines Prozentes aller Männer, die ihre Frau schlagen, diese auch umbringen. Er brachte dies in Verteidigung seines Mandanten vor, der die ermordete Nicole Brown Simpson erwiesenermassen geschlagen hatte. Dies wurde von den Strafverfolgungsbehörden und den Medien als belastendes Indiz gegen O.J. Simpson gewertet.
- 413 Das Argument von DERSHOWITZ ist auf den ersten Blick einleuchtend: es ist sehr, sehr selten, dass ein Mann, der seine Frau schlägt, diese auch umbringt, daher ist das Indiz, dass O.J. Simpson seine Frau geschlagen hat, auch nicht belastend. „Ermorden“ ist mit anderen Worten nicht repräsentativ, oder typisch, für „schlagen“.
- 414 Der Statistiker IRVING J. GOOD wies in einem Leserbrief an die Zeitschrift *Nature* nach, dass die Argumentation von DERSHOWITZ nicht richtig ist.⁶⁷³

⁶⁶⁹ D.h. es wird eine Übereinstimmung der Proben gefunden, obwohl keine vorliegt.

⁶⁷⁰ WILLIAM C. THOMPSON/FRANCO TARONI/COLIN C.G. AITKEN, How the Probability of a False Positive Affects the Value of DNA Evidence, *Journal of Forensic Science* 2003, 47-54, 49.

⁶⁷¹ Siehe THOMPSON/TARONI/AITKEN, FN 670, 49 ff. für eine Erklärung.

⁶⁷² ZR 94 (1995) Nr. 7, 26.

⁶⁷³ IRVING J. GOOD, When Batterer Turns Murderer, *Nature* 1995, 541.

- 415 Angenommen, E bezeichnet „ermordet (egal von wem) in 1994“, S „schuldig“ (d.h. der Mann der geschlagenen Frau ist der Mörder), $\neg S$ „unschuldig“ und G „geschlagen von ihrem Mann“. Dann gilt für das Verhältnis der bedingten Wahrscheinlichkeiten gemäss dem Bayes-Theorem in seiner Chancen Form (siehe Gleichung (7))

$$\frac{P(S | E \text{ und } G)}{P(\neg S | E \text{ und } G)} = \frac{P(S | G)}{P(\neg S | G)} \cdot \frac{P(E | S \text{ und } G)}{P(E | \neg S \text{ und } G)} \quad (13)$$

- 416 GOOD nimmt die Wahrscheinlichkeit von DERSHOWITZ, dass ein seine Frau schlagender Mann diese umbringt (1/1'000), und multipliziert sie mit 1/10, um die Wahrscheinlichkeit anzugeben, dass der Mord in einem bestimmten Jahr geschieht.⁶⁷⁴ Die Anfangswahrscheinlichkeit, dass ein schlagender Mann seine Frau in einem bestimmten Jahr umbringt, ist daher

$$P(S | G) = 1/10'000 \text{ (0,0001)} \quad (14)$$

- 417 Die Wahrscheinlichkeit, dass eine Frau ermordet wurde, dann wenn ihr Mann sie ermordet hat, ist logischerweise 1, also

$$P(E | S \text{ und } G) = P(E | S) = 1 \quad (15)$$

- 418 GOOD stellt weiter fest, dass in den USA jährlich rund 25'000 Personen ermordet werden; bei einer Bevölkerung von rund 250 Millionen beträgt daher die Wahrscheinlichkeit für eine beliebige Person,⁶⁷⁵ und somit auch für die Frau eines schlagenden Mannes, in einem bestimmten Jahr ermordet zu werden

$$P(E | \neg S \text{ und } G) = P(E | \neg S) \approx 1/10'000 \text{ (0,0001)} \quad (16)$$

- 419 Setzt man nun die Zahlen in die Gleichung (7) ein, so erhält man

$$\frac{P(S | E \text{ und } G)}{P(\neg S | E \text{ und } G)} = \frac{0,0001}{0,9999} \cdot \frac{1}{0,0001} \approx 1 \quad (17)$$

- 420 Mit anderen Worten beträgt die Wahrscheinlichkeit, dass die getötete Frau eines Mannes, der sie schlägt, von ihrem Mann ermordet wurde, rund 50 % ($P(S | E \text{ und } G) \approx 0,5$). Es ist daher nur vernünftig, den Ehemann als Hauptverdächtigen anzusehen, wenn feststeht, dass

⁶⁷⁴ GOOD geht offensichtlich davon aus, dass ein Mann, der seine Frau umbringt, im Schnitt zehn Jahre mit ihr verheiratet ist. Er berechnet jedoch auch noch eine Variante, in der der Mann im Schnitt 20 Jahre verheiratet ist, ehe er seine Frau umbringt. Die Wahrscheinlichkeit eines Mordes in einem bestimmten Jahr ist entsprechend geringer, aber das Endresultat immer noch sehr belastend für O.J.

⁶⁷⁵ Natürlich ist das Risiko, ermordet zu werden, nicht gleichmässig verteilt. Für junge Schwarze in bestimmten Stadtteilen von Los Angeles ist es erheblich höher als für weisse Frauen reicher Sportstars. Die Annahme der gleichmässigen Verteilung ist aber hier angezeigt, weil keine spezifischeren Daten vorhanden sind und wirkt sich zugunsten des Verdächtigen aus.

er seine Frau geschlagen hat und diese tot aufgefunden wird. Dass „ermorden“ untypisch ist für schlagende Männer, ändert nichts daran.⁶⁷⁶

- 421 GERD GIGERENZER weist darauf hin, dass die obige – gerade für Juristen – einigermaßen komplexe Rechnung erheblich vereinfacht werden kann, wenn man sie in natürlichen Häufigkeiten darstellt.⁶⁷⁷ Man stelle sich 10'000 geschlagene Frauen vor. Von diesen wird in einem bestimmten Jahr eine von ihrem Ehemann ermordet. Von den restlichen 9'999 wird statistisch gesehen eine von jemand anderem ermordet. Von den zwei ermordeten Frauen wurde also eine von ihrem schlagenden Ehemann ermordet; mit anderen Worten beträgt die Wahrscheinlichkeit, dass der schlagende Ehemann der Mörder ist, 50 %.

B. Kritik am deskriptiven Modell in der Psychologie

- 422 Die Forschung von KAHNEMAN und TVERSKY zur Repräsentativitätsheuristik, und insbesondere zum *base rate neglect*, ist in den neunziger Jahren des letzten Jahrhunderts zunehmend kritisiert worden. GERD GIGERENZER,⁶⁷⁸ JONATHAN J. KOEHLER⁶⁷⁹ und LEDA COSMIDES/JOHN TOOBY⁶⁸⁰ kritisieren, dass die von KAHNEMAN und TVERSKY aufgrund ihrer Experimente gezogenen weitreichenden Schlüsse nicht zuträfen und nicht generell davon gesprochen werden könne, dass Menschen intuitiv schlechte Statistiker seien.
- 423 Der wohl schärfste Kritiker GIGERENZER bringt zwei Hauptargumente vor, eines logisch-mathematischer Art, das andere empirischer Natur. Erstens weist er darauf hin, dass umstritten ist, was es bedeutet, dass ein Ereignis „wahrscheinlich“ ist, und dass man sinnvollerweise gar nicht von der Wahrscheinlichkeit eines singulären Ereignisses sprechen könne. Es gebe daher keine normative Basis, mit der man die von den Versuchspersonen gegebenen Antworten zur Wahrscheinlichkeit eines singulären Ereignisses vergleichen und die Abweichungen als „Fehler“ qualifizieren könne. Zweitens, so das empirische Argument, verschwänden die beobachteten Phänomene weitgehend, wenn die Information in der Form natürlicher Häufigkeiten präsentiert würden. Die Kritik von GIGERENZER war mächtig genug, um KAHNEMAN und TVERSKY zu einer Antwort⁶⁸¹ herauszufordern, obwohl sie normalerweise bewusst darauf verzichtet haben, explizit zur Kritik an ihrer Arbeit Stellung

⁶⁷⁶ Ähnlich wie DERSHOWITZ, aber in einem anderen Fall, argumentieren DEBORAH DAVIS/WILLIAM C. FOLLETTE, Rethinking the Probative Value of Evidence: Base Rates, Intuitive Profiling, and the „Post-diction“ of Behavior, Law and Human Behavior 2002, 133-158; sie werden durch RICHARD D. FRIEDMANN/ROGER C. PARK, Sometimes What Everybody Thinks They Know is True, Law and Human Behavior 2003, 629-644, die ähnlich wie GOOD argumentieren, und DAVID H. KAYE/JONATHAN J. KOEHLER, The Misquantification of Probative Value, Law and Human Behavior 2003, 645-659, die darauf hinweisen, dass sich die abstrakte Beweiskraft eines Indizes im Likelihood-Quotienten spiegelt, überzeugend widerlegt.

⁶⁷⁷ GERD GIGERENZER, Ecological Intelligence: An Adaptation for Frequencies, in: DENISE D. CUMMINS/COLIN ALLEN (Hrsg.), The Evolution of Mind, New York 1998, 9-29, 19.

⁶⁷⁸ GERD GIGERENZER, FN 222; ders., Why the Distinction between Single-event Probabilities and Frequencies is Important for Psychology (and Vice Versa), in: WRIGHT/AYTON (Hrsg.), 129-161.

⁶⁷⁹ KOEHLER, FN 633, 1 ff.

⁶⁸⁰ LEDA COSMIDES/JOHN TOOBY, Are Humans Good Intuitive Statisticians After All? Rethinking Some Conclusions from the Literature on Judgment Under Uncertainty, Cognition 1996, 1-73.

⁶⁸¹ KAHNEMAN/TVERSKY, FN 179.

zu nehmen.⁶⁸² Im Folgenden werden die Argumente von GIGERENZER und die sich daraus ergebenden Schlussfolgerungen dargestellt.

1. Was heisst „ein Ereignis tritt mit einer Wahrscheinlichkeit von 0,x auf“?

- 424 Was bedeutet die Aussage „Eine Frau im gebärfähigen Alter, die mit einem zeugungsfähigen Mann Geschlechtsverkehr hat, wird mit einer Wahrscheinlichkeit von 0,05 schwanger“? Um sagen zu können, was eine Aussage bedeutet, muss man aufzeigen, was der Fall sein muss, damit die Aussage wahr ist. Im Fall der Wahrscheinlichkeit der Schwangerschaft lässt sich durch Beobachtung eines einzelnen Ereignisses nicht feststellen, ob die Aussage zutrifft. Entweder wird die Frau schwanger, dann ist die Wahrscheinlichkeit 1, oder sie wird nicht schwanger, dann ist die Wahrscheinlichkeit Null. Ein bisschen schwanger kann man bekanntlich nicht werden. So oder anders wird die Aussage, dass sie mit einer „Wahrscheinlichkeit von 0,05“ schwanger wird, durch die Beobachtung der Folgen des Geschlechtsverkehrs weder bestätigt noch widerlegt. Ob die Aussage zutrifft, muss daher an einem anderen Kriterium gemessen werden.

a. Objektive Wahrscheinlichkeit

- 425 Eine Möglichkeit, Wahrscheinlichkeit zu definieren, ist als *Grenzwert der relativen Häufigkeit* eines Merkmals in einem Kollektiv (*long run frequency*).⁶⁸³ Nach dieser Auffassung kann man von Wahrscheinlichkeit nur sprechen, wenn „es sich um eine in grossen Mengen auftretende Erscheinung, [...] um eine praktisch unbegrenzte Folge von gleichartigen Beobachtungen handelt“.⁶⁸⁴ Dieser Wahrscheinlichkeitsbegriff wird wahlweise *objektiv*, *statistisch*, *empirisch* oder *frequentistisch* genannt. Von der Wahrscheinlichkeit eines singulären Ereignisses zu sprechen ist nach dieser Auffassung sinnlos, weil es für ein singuläres Ereignis keine Referenzklasse gibt, in der man die Häufigkeit des Vorkommens des Ereignisses zählen könnte.⁶⁸⁵ Der Anwendungsbereich der objektiven Wahrscheinlichkeitstheorie ist deshalb auf Gebiete begrenzt, in denen lange Beobachtungsreihen gemacht werden können, wie in der Physik, bei Glücksspielen oder im Versicherungswesen. Nach der objektiven Wahrscheinlichkeitstheorie kann die obige Aussage zur Schwangerschaft nur so verstanden werden, dass von 1'000 Frauen im gebärfähigen Alter, die mit tausend zeugungsfähigen Männern Geschlechtsverkehr haben, 50 schwanger werden. Daraus abzuleiten, dass die 1'001. Frau mit einer Wahrscheinlichkeit von 0,05 schwanger wird, ist unzulässig.

⁶⁸² DANIEL KAHNEMAN, *Autobiography*, 2002 (nobelprize.org/economics/laureates/2002/kahneman-autobio.html; besucht am 12. Januar 2005).

⁶⁸³ RICHARD VON MISES, *Wahrscheinlichkeit, Statistik und Wahrheit*, 4. Aufl., Wien/New York 1972, 263.

⁶⁸⁴ VON MISES, FN 683, 12.

⁶⁸⁵ VON MISES, FN 683, 10.

b. Subjektive Wahrscheinlichkeit

- 426 Nach dem *subjektiven* oder *epistemischen* Wahrscheinlichkeitsbegriff ist Wahrscheinlichkeit Ausdruck unseres subjektiven Glaubens an das Eintreten oder die Ursache eines Ereignisses in Anbetracht der derzeit zur Verfügung stehenden Informationen. Der Nachsatz „in Anbetracht der derzeit zur Verfügung stehenden Informationen“ ist wichtig; nach der subjektiven Wahrscheinlichkeitstheorie ist es sinnlos, von der Wahrscheinlichkeit eines Ereignisses „an sich“ zu sprechen. Die Wahrscheinlichkeit eines Ereignisses hängt immer von den bekannten Informationen ab.⁶⁸⁶ Ändern sich diese, muss die Wahrscheinlichkeit unter Berücksichtigung der neuen Informationen neu berechnet werden. Die frühere Wahrscheinlichkeit ist deswegen nicht falsch, sondern ging eben von einem anderen Wissensstand aus. So lässt sich auch zwanglos erklären, dass ein Wahrscheinlichkeitsurteil desto besser (begründet) ist, je mehr Informationen es berücksichtigt: mit der Verbreiterung der Informationsbasis wird es von einer unsicheren zu einer „sichereren“ Entscheidungsgrundlage.⁶⁸⁷ Im obigen Beispiel mit der Wahrscheinlichkeit der Schwangerschaft wird sich subjektive Wahrscheinlichkeit beispielsweise ändern, wenn man erfährt, dass die Frau im Zeitpunkt des Geschlechtsverkehrs ihren Eisprung hat. Vielleicht würde man statt von „subjektiver Wahrscheinlichkeit“ besser von „Gewissheit“ oder „Überzeugungsgrad“ sprechen, weil subjektive Wahrscheinlichkeit letztlich einen Grad von Gewissheit ausdrückt.⁶⁸⁸
- 427 Nach dem subjektiven Wahrscheinlichkeitsbegriff lässt sich die Aussage, dass es morgen mit einer Wahrscheinlichkeit von 0,3 regnen wird, in eine Wette mit einer Chance von 3:7 übersetzen. D.h. die Person, die die Aussage macht, müsste bereit sein, eine Wette einzugehen, bei der sie bei einem Einsatz von 3 Einheiten 10 Einheiten gewinnen kann (und somit einen Gewinn von 7 macht), wenn es morgen regnet.⁶⁸⁹ Dieser von FRANK P. RAMSEY entwickelte Gedanke erlaubt, Sätze der subjektiven Wahrscheinlichkeitstheorie auf einleuchtende Weise zu beweisen; ein Satz ist dann richtig, wenn er keine Wette zulässt, bei der der Wettgegner in jedem Fall verliert (sog. „*Dutch book*“).⁶⁹⁰ Ob die Theorie rationalen Wettens geeignet ist, die Überzeugung eines Richters festzustellen, bleibe dahingestellt,⁶⁹¹ sie ist auf jeden Fall nicht notwendig, um den subjektiven Wahrscheinlichkeitsbegriff anzuwenden.⁶⁹²

⁶⁸⁶ DENNIS LINDLEY, Foundations, in: AYTON/WRIGHT, FN 619, 3-15, 6.

⁶⁸⁷ ERNST LUDWIG NELL, Wahrscheinlichkeitsurteile in juristischen Entscheidungen, Diss. Bayreuth 1982, Berlin 1983, 58.

⁶⁸⁸ MINNA GRÄNS, Das Risiko materiell fehlerhafter Urteile, Diss. Uppsala 1995, München 2002, 157.

⁶⁸⁹ CHRISTOPHER HITCHCOCK, Probability and Chance: Philosophical Aspects, in: SMELSER/BALTES, FN 652, 12089-12095, 12091.

⁶⁹⁰ JEFFREY, FN 613, 5 ff.

⁶⁹¹ Zweifelnd BENDER/NACK, FN 605, Rz. 369.

⁶⁹² ELEONORA BOURMISTROV-JÜTTNER, Subjektive Wahrscheinlichkeitstheorie und rationale Entscheidungstheorie in Anwendung auf die Rechtspraxis, Diss. München 1987, 48 f.

- 428 Der subjektive Wahrscheinlichkeitsbegriff wurde im 20. Jahrhundert von FRANK P. RAMSEY,⁶⁹³ BRUNO DE FINETTI⁶⁹⁴ und LEONHARD SAVAGE⁶⁹⁵ entwickelt. Die subjektive Wahrscheinlichkeitstheorie befasst sich naturgemäss in erster Linie mit dem Problem, wie die Wahrscheinlichkeit eines Ereignisses in Anbetracht neuer Informationen zu revidieren ist. Sie stellt Regeln auf, wie Überzeugungen rational und *kohärent*, d.h. widerspruchsfrei, verknüpft werden können. Da das Bayes-Theorem genau dies zeigt, wird die subjektive Wahrscheinlichkeitstheorie oft auch mit „Bayes’scher Statistik“ (*Bayesian statistics*) bezeichnet.
- 429 Zwei Menschen, die den gleichen Informationsstand aufweisen, können einem Ereignis durchaus eine andere subjektive Wahrscheinlichkeit beimessen, auch wenn es in der Praxis Fälle geben wird, in denen die meisten Leute sich einig sind.⁶⁹⁶ Genau dies haben KAHNEMAN und TVERSKY stillschweigend angenommen, und hier setzt die (eine) Kritik von GIGERENZER an.⁶⁹⁷ KAHNEMAN und TVERSKY setzen voraus, dass beim Anwalt/Ingenieur Problem (das auf S. 129 vorgestellt wurde) die Anfangswahrscheinlichkeit, dass es sich um einen Ingenieur handelt, 0,3 beträgt, wenn die Stichprobe 30 Ingenieure und 70 Anwälte enthält. Aber die subjektive Wahrscheinlichkeitstheorie verlangt nicht, dass die relative Häufigkeit der Verteilung mit der Anfangswahrscheinlichkeit gleichzusetzen ist. Die Anfangswahrscheinlichkeit lässt sich nach der subjektiven Wahrscheinlichkeitstheorie nicht bestimmen, sie muss als gegeben hingenommen werden. Weil aber die Belastungswahrscheinlichkeit von der Anfangswahrscheinlichkeit abhängt, lassen sich die Antworten der Versuchspersonen nach der subjektiven Wahrscheinlichkeitstheorie auch nicht als richtig oder falsch qualifizieren, wenn man deren subjektive Anfangswahrscheinlichkeit nicht kennt.⁶⁹⁸

c. Wahrscheinlichkeitsbegriff und Tatsachenfeststellung vor Gericht

- 430 Die Diskussion, welcher Wahrscheinlichkeitsbegriff der „richtige“ ist, wird mit Sicherheit auch im 21. Jahrhundert weitergeführt werden.⁶⁹⁹ In der deutschen juristischen Literatur vertreten BENDER/NACK,⁷⁰⁰ ERNST LUDWIG NELL,⁷⁰¹ ELEONORA BOURMISTROV-JÜTTNER⁷⁰² und MINNA GRÄNS⁷⁰³ einen subjektiven Wahrscheinlichkeitsbegriff, während HERMANN WEITNAUER,⁷⁰⁴ RUPERT SCHREIBER,⁷⁰⁵ HANS JOACHIM MUSIELAK⁷⁰⁶ und BERNHARD U.

⁶⁹³ FRANK P. RAMSEY, *Truth and Probability*, London 1926, in: RICHARD B. BRAITHWAITE (Hrsg.), *The Foundations of Mathematics and other Logical Essays*, New York 1931, 156-198 (elektronische Version erhältlich unter cepa.newschool.edu/het/texts/ramsey/ramsess.pdf (besucht am 12. Januar 2005)).

⁶⁹⁴ BRUNO DE FINETTI, *Theory of Probability, a Critical Introductory Treatment*, Chichester 1974.

⁶⁹⁵ SAVAGE, FN 316.

⁶⁹⁶ LINDLEY, FN 686, 6.

⁶⁹⁷ GIGERENZER, FN 222, 145.

⁶⁹⁸ COSMIDES/TOBY, FN 680, 5.

⁶⁹⁹ HITCHCOCK, FN 689, 12095.

⁷⁰⁰ BENDER/NACK, FN 605, Rz. 368.

⁷⁰¹ NELL, FN 687, 59 f.

⁷⁰² BOURMISTROV-JÜTTNER, 692, 62 ff.

⁷⁰³ GRÄNS, FN 688, 159.

⁷⁰⁴ HERMANN WEITNAUER, *Wahrscheinlichkeit und Tatsachenfeststellung*, in *Karlsruher Forum* 1966, 6.

MAASSEN⁷⁰⁷ den objektiven Wahrscheinlichkeitsbegriff im Sinne der relativen Häufigkeit vertreten.⁷⁰⁸ REINHARD GREGER lehnt sowohl eine Anwendung des subjektiven wie des objektiven Wahrscheinlichkeitsbegriffs im Rahmen der Tatsachenfeststellung vor Gericht ab, da sich seiner Auffassung nach juristische Urteile nicht mit wahrscheinlichkeitstheoretischen Erwägungen erfassen lassen.⁷⁰⁹

- 431 In der Schweiz hat sich in der Fachliteratur noch niemand dazu geäußert, ob bei Tatsachenfeststellungen vor Gericht von einem objektiven oder subjektiven Wahrscheinlichkeitsbegriff ausgegangen werden muss. CHRISTOPHE CHAMPOD und FRANCO TARONI verwenden einen objektiven Wahrscheinlichkeitsbegriff, befassen sich aber schwerewichtig mit Beweismitteln der forensischen Wissenschaften, wo objektive Wahrscheinlichkeiten empirisch feststellbar sind.⁷¹⁰ Meines Erachtens gibt es gute Gründe dafür, anzunehmen, dass das Rechtssystem sich für den subjektiven Wahrscheinlichkeitsbegriff entschieden hat; schon deshalb, weil es kaum identische Beweissituationen gibt, die sich beliebig oft wiederholen liessen.⁷¹¹ Der Richter beurteilt immer den Einzelfall, und die Wahrscheinlichkeit eines Einzelfalls lässt sich nur in subjektiver Wahrscheinlichkeit ausdrücken.⁷¹²
- 432 Die schweizerische Lehre und Rechtsprechung umschreibt das Beweismass des Vollbeweises mit den folgenden Formulierungen: „Die richterliche Überzeugung braucht [...] keine Gewissheit zu sein. Mit Gewissheit lassen sich Tatsachen, insbesondere solche der Vergangenheit oder der Zukunft, praktisch kaum feststellen. Die richterliche Überzeugung darf sich auf eine Wahrscheinlichkeit stützen, die zwar objektiv den Zweifel nicht ausschliesst, ihn aber subjektiv, nach Ansicht des Gerichts, als unbegründet erscheinen lässt. [...] Der Richter braucht nicht überzeugt zu sein, dass seine Annahme unter allen Umständen richtig sein müsse, es genügt, wenn er zu dem Ergebnis gelangt, dass sie nach den konkreten Umständen die am ehesten vertretbare sei“.⁷¹³ „Absolute Gewissheit gibt es nicht und kann es nicht geben, und der Prozess kann sie nicht fordern. Im Rechtssinn ist daher ein Beweis bereits erbracht, wenn der Richter sich überzeugt hat, wenn also ein so hoher Grad von Wahrscheinlichkeit besteht, dass vernünftigerweise mit der Möglichkeit des Gegenteils nicht mehr zu rechnen ist“.⁷¹⁴ „Tatbestandsmerkmale müssen im Allgemeinen bewiesen, d.h. zur vollen Überzeugung des Richters gebracht werden. Hiefür kann es

⁷⁰⁵ RUPERT SCHREIBER, *Theorie des Beweiswerts für Beweismittel im Zivilprozessrecht*, Berlin etc. 1968.

⁷⁰⁶ HANS JOACHIM MUSIELAK, *Die Grundlagen der Beweislast im Zivilprozessrecht*, Berlin/New York 1978, 112.

⁷⁰⁷ BERNHARD U. MAASSEN, *Beweismassprobleme im Schadenersatzprozess*, Köln etc. 1975, 6.

⁷⁰⁸ Die Hinweise verdanke ich GRÄNS, FN 688, 158.

⁷⁰⁹ REINHARD GREGER, *Beweis und Wahrscheinlichkeit*, Diss. Erlangen-Nürnberg, München 1978, 49, 55.

⁷¹⁰ CHRISTOPHE CHAMPOD/FRANCO TARONI, *Probabilités au procès pénal – risques et solutions*, Schweizerische Zeitschrift für Strafrecht 1994, 195-219, 215.

⁷¹¹ GRÄNS, FN 688, 159.

⁷¹² NELL, FN 687, 59; PETER DONNELLY/RICHARD D. FRIEDMAN, *DNA Database Searches and the Legal Consumption of Scientific Evidence*, Michigan Law Review 1997, 931-984, 970: “Thus, the legal system embraces subjectivity, for it has no other choice”.

⁷¹³ MAX GULDENER, *Beweiswürdigung und Beweislast nach schweizerischem Zivilprozessrecht*, Zürich 1955, 6 f.

⁷¹⁴ MAX KUMMER, *Grundriss des Zivilprozessrechts*, 4. Aufl., Bern 1984, 135.

aber genügen, dass ein Sachverhalt mit an Sicherheit grenzender Wahrscheinlichkeit erstellt ist, wenn völlige Sicherheit nicht zu gewinnen ist“.⁷¹⁵ „Ein Beweis gilt als erbracht, wenn das Gericht nach objektiven Gesichtspunkten von der Richtigkeit einer Sachbehauptung überzeugt ist. Absolute Gewissheit kann dabei nicht verlangt werden. Es genügt, wenn das Gericht am Vorliegen der behaupteten Tatsache keine ernsthaften Zweifel mehr hat oder allenfalls verbleibende Zweifel als leicht erscheinen“ (BGE 130 II 321, 325).⁷¹⁶

433 Diese Formulierungen stellen einerseits klar, dass nicht Sicherheit verlangt wird. Der Richter darf auch dann den Beweis als erbracht ansehen, wenn eine kleine Wahrscheinlichkeit bleibt, dass seine Überzeugung nicht zutrifft. Obwohl in den Beispielen meist von „subjektiver Gewissheit“ und „objektiver Wahrscheinlichkeit“ gesprochen wird, darf man dies nicht dahingehend verstehen, dass damit einem objektiven Wahrscheinlichkeitsbegriff das Wort geredet würde. Die Juristen, die die obigen Regeln formuliert haben, haben sich nicht mit Wahrscheinlichkeitstheorie befasst und waren sich der unterschiedlichen Wahrscheinlichkeitsbegriffe kaum bewusst. Man kann natürlich argumentieren, dass der Begriff der Wahrscheinlichkeit, der in der Jurisprudenz verwendet wird, weder dem objektiven noch dem subjektiven Wahrscheinlichkeitsbegriff entspricht, sondern etwas Eigenes ist. Damit entzieht man juristische Urteile weitestgehend der rationalen Kontrolle.⁷¹⁷

434 Die Umschreibungen der herrschenden Lehre lassen sich aber zwanglos als Umschreibungen der subjektiven Wahrscheinlichkeit, d.h. des verlangten Grades der Überzeugung des Richters, verstehen⁷¹⁸ (dies im Übrigen auch dann, wenn man der so genannten „objektiven Beweismasstheorie“ folgt; auch diese begnügt sich mit Wahrscheinlichkeit).⁷¹⁹ BOURMISTROV-JÜTTNER betont:⁷²⁰

So verstanden drücken die präzisierten juristischen Wahrscheinlichkeitsurteile nichts anderes aus als die rationale Überzeugung des Richters vom (oder den rationalen Glauben des Richters an das) Bestehen bestimmter Sachverhalte aufgrund der zur Verfügung stehenden Informationen und Erfahrungen.

435 Das Verständnis des juristischen Wahrscheinlichkeitsbegriffs als subjektive Wahrscheinlichkeit hat den Vorteil, dass die Kohärenz und Rationalität juristischer Urteile am normativen Standard der subjektiven Wahrscheinlichkeitstheorie gemessen und intersubjektiv diskutiert werden können.⁷²¹

436 Aussagen wie „Je unwahrscheinlicher die Tatsachenbehauptung einer Partei ist, desto höhere Anforderungen darf das Gericht an den Beweis der behaupteten Tatsachen stellen“⁷²²

⁷¹⁵ OSCAR VOGEL/KARL SPÜHLER, Grundriss des Zivilprozessrechts, 7. Aufl., Bern 2001, 256.

⁷¹⁶ Die schweizerischen Prozessgesetze verzichten in der Regel auf eine Umschreibung des Beweismasses und begnügen sich damit, festzuhalten, dass der Richter die Beweise frei würdigt; Art. 40 BZPO, § 148 ZPO-ZH, Art. 219 ZPO-BE.

⁷¹⁷ BOURMISTROV-JÜTTNER, FN , 78, spricht von „erkenntnistheoretischem Esoterismus“.

⁷¹⁸ So auch für das deutsche Recht BENDER/NACK, FN 605, 368.

⁷¹⁹ GRÄNS, FN 688, 175.

⁷²⁰ BOURMISTROV-JÜTTNER, FN 692, 61.

⁷²¹ BOURMISTROV-JÜTTNER, FN 692, 78.

⁷²² MAX GULDENER, Schweizerisches Zivilprozessrecht, 3. Aufl., Zürich 1979, 322; zustimmend RICHARD FRANK/HANS STRÄULI/GEORG MESSMER, Kommentar zur zürcherischen Zivilprozessordnung, 3. Aufl. Zürich 1997, § 148 N 8.

lassen sich durch das Bayes-Theorem leicht erklären. Für das Bundesgericht gehört diese Überlegung in den Bereich der Beweiswürdigung und darf nicht zu einem erhöhten Beweismass führen (BGE 130 II 321, 325). Dort gehört sie nach dem Bayes-Theorem auch hin. Sie besagt nichts anderes, als dass, wenn eine geringe Anfangswahrscheinlichkeit des Vorliegens einer Tatsache besteht, die abstrakte Beweiskraft der Beweismittel höher sein muss, ehe das feststehende Beweismass, die subjektive Überzeugung des Richters vom Vorliegen der Tatsache, erreicht wird.

- 437 Das Recht verlangt auch, dass das Gericht seine freie Beweiswürdigung nachvollziehbar begründet. Freie Beweiswürdigung „bedeutet nicht Willkür, sondern Pflicht zu gewissenhafter Schlussfolgerung aufgrund des Ergebnisses des Beweisverfahrens. Das Gericht muss seine Ansicht begründen können, es genügt nicht, wenn es sich einfach für überzeugt erklärt“. ⁷²³ Die Beweiswürdigung „darf die Gesetze der Logik nicht verletzen“, ⁷²⁴ der Richter muss „das Beweismaterial sorgfältig prüfend bewerten und begründen können, weswegen er hier den Beweis als erbracht und dort als fehlgeschlagen erachtet“. ⁷²⁵ Die richterliche Überzeugung ist in den Urteilerwägungen zu begründen, damit die richterliche Überzeugungsbildung verstandesmässig kontrolliert werden kann. ⁷²⁶
- 438 Dort, wo objektive Wahrscheinlichkeiten empirisch feststellbar sind, darf der Richter daher meines Erachtens nur diese objektiven Wahrscheinlichkeiten zu seiner subjektiven Anfangswahrscheinlichkeit machen. ⁷²⁷ Dies ist letztlich auch die Auffassung von WEITNAUER und (ihm folgend) MUSIELAK, die die Anwendung des objektiven Wahrscheinlichkeitsbegriffs mit dem Argument verteidigen, beim Einzelfall handle es sich um eine „Stichprobe“ aus dem Kollektiv. Die Aussage, dass in einer Gesamtmenge von 100 Personen 90 ein bestimmtes Merkmal aufweisen, lässt sich nach ihrer Auffassung auf den Einzelfall übertragen, indem die Richtigkeit der Hypothese, dass die konkrete Person das betreffende Merkmal aufweist, mit einer Wahrscheinlichkeit von 90 % zutrefte, wenn die Person zufällig aus dem Kollektiv ausgewählt wurde. ⁷²⁸ Dies ist keine Folgerung aus der Wahrscheinlichkeitstheorie, sondern ein Werturteil des Rechtssystems, das intersubjektiv diskutierbare Entscheidungsbegründungen verlangt. Beim Anwalt/Ingenieur Problem muss der Richter daher von der empirischen Wahrscheinlichkeit von 0,3 ausgehen, dass es sich bei der beschriebenen Person um einen Ingenieur handelt. Dies ist zwar nicht logisch zwingend, aber jede Abweichung davon ist zumindest erklärungsbedürftig. Das Rechtssystem hat sich daher, wenn man so will, für die pragmatische, wenn auch logisch nicht begründbare, Auffassung von KAHNEMAN und TVERSKY entschieden, dass man bei empirisch feststellbaren Wahrscheinlichkeiten diese als subjektive a priori Wahrscheinlichkeiten annehmen und *unbegründete* Abweichungen von der so angenommenen Anfangswahrscheinlichkeit als Fehler betrachten muss.

⁷²³ GULDENER, FN 722, 322.

⁷²⁴ WALTER HABSCHIED, Schweizerisches Zivilprozess- und Gerichtsorganisationsrecht, 2. Aufl. Zürich 1990, 398.

⁷²⁵ KUMMER, FN 714, 136.

⁷²⁶ GULDENER, FN 713, 6.

⁷²⁷ Ähnlich BENDER/NACK, FN 605, 369.

⁷²⁸ WEITNAUER, FN 704, 20 und 46; MUSIELAK, FN 706, 112.

2. Häufigkeiten statt Wahrscheinlichkeiten

- 439 GIGERENZER und COSMIDES/TOOBY behaupten, dass die beobachteten Verletzungen des Bayes-Theorem weitgehend verschwinden, wenn man den Versuchspersonen die Informationen in der Form von natürlichen Häufigkeiten („5 von 20“) statt als Wahrscheinlichkeiten (0,25) präsentiert.

Information	Darstellung der Information
<i>Natürliche Häufigkeiten</i>	
$P(H)$	100 von 10'000 getesteten Autofahrern sind betrunken.
$P(I H)$	Bei 95 von 100 betrunkenen Autofahrern verfärbt sich das Teströhrchen.
$P(I \neg H)$	Bei 495 von 9'900 nüchternen Autofahrern verfärbt sich das Teströhrchen auch.
<i>Normalisierte Häufigkeiten</i>	
$P(H)$	100 von 10'000 getesteten Autofahrern sind betrunken.
$P(I H)$	Bei 9'500 von 10'000 betrunkenen Autofahrern verfärbt sich das Teströhrchen.
$P(I \neg H)$	Bei 500 von 10'000 nüchternen Autofahrern verfärbt sich das Teströhrchen auch.
<i>Relative Häufigkeiten</i>	
$P(H)$	1 % aller getesteten Autofahrer ist betrunken.
$P(I H)$	Bei 95 % aller betrunkenen Autofahrer verfärbt sich das Teströhrchen.
$P(I \neg H)$	Bei 5 % aller nüchternen Autofahrer verfärbt sich das Teströhrchen auch.
<i>Wahrscheinlichkeiten</i>	
$P(H)$	Die Wahrscheinlichkeit, dass ein getesteter Autofahrer betrunken ist, beträgt 0,1.
$P(I H)$	Die Wahrscheinlichkeit, dass sich das Teströhrchen bei einem betrunkenen Autofahrer verfärbt, beträgt 0,95.
$P(I \neg H)$	Die Wahrscheinlichkeit, dass sich das Teströhrchen bei einem nüchternen Autofahrer auch verfärbt, beträgt 0,05.

Tabelle 10: Verschiedene Darstellungsformen der Information im „Alkotest“ Problem (S. 161)

- 440 GIGERENZER stellt fest, dass sich die *heuristics and biases* Literatur fast ausschliesslich auf die Algorithmen der Informationsverarbeitung konzentriert und dabei die Darstellung der Information und deren Einfluss auf die Informationsverarbeitung vernachlässigt hat.⁷²⁹ Information braucht Repräsentation. Wenn unser Gehirn für eine Art der Darstellung von Wahrscheinlichkeits-Informationen besser gerüstet ist, dann für die Darstellung der Information in der Form natürlicher Häufigkeiten. Prozente und Brüche waren erst nach der Entwicklung komplexer mathematischer Systeme verständlich. Natürliche Häufigkeiten aber konnten schon vor der Erfindung von Schrift und Mathematik beobachtet werden.⁷³⁰ Unsere prähistorischen Vorfahren konnten feststellen, dass sie bei der Jagd im nördlichen Wald fünf von den letzten zwanzig Malen erfolgreich waren; sie haben sich kaum ausgerechnet, dass dies einer Wahrscheinlichkeit von 0,25 entspricht, zumal bei der Umrechnung von natürlichen Häufigkeiten in Wahrscheinlichkeiten wichtige Informationen, beispielsweise über die Anzahl der Beobachtungen, auf der eine Wahrscheinlichkeit beruht,

⁷²⁹ GIGERENZER, FN 222, 141.

⁷³⁰ GIGERENZER, FN 222, 142.

verloren gehen.⁷³¹ Am besten sollte nach dieser Theorie unsere Wahrscheinlichkeitsverarbeitung sein, wenn sie auf der durch eigene Beobachtung festgestellten natürlichen Häufigkeit eines Ereignisses (*natural sampling*) beruht.⁷³²

- 441 GIGERENZER und COSMIDES/TOOBY berichten übereinstimmend, dass die beobachteten Abweichungen vom Bayes-Theorem weitgehend verschwinden, wenn man die klassischen Fragen aus der *heuristics and biases* Literatur umformuliert und natürliche Häufigkeiten statt Wahrscheinlichkeiten verwendet. So steigt die Anzahl der richtigen Antworten im „medizinischen Diagnose-Problem“⁷³³ (sehr ähnlich zum hinten dargestellten Alko-Test Problem) von 12 % bei der Darstellung der Information in Wahrscheinlichkeiten auf 46 %⁷³⁴, 56 %⁷³⁵ oder 73 %⁷³⁶ bei der Darstellung in natürlichen Häufigkeiten.
- 442 KAHNEMAN und TVERSKY weisen in ihrer Antwort auf die Kritik erstens darauf hin, dass sie nie behauptet hätten, die Repräsentativitätsheuristik trete unter allen Umständen auf. Natürlich gebe es Umstände, in denen Menschen eher den Regeln der Wahrscheinlichkeitstheorie folgen würden. Wenn GIGERENZER nun nachweise, dass Menschen nicht immer die von ihnen prognostizierten Fehler machten, habe er sie keineswegs widerlegt.⁷³⁷ Die Schlussfolgerungen von GIGERENZER seien zudem empirisch ungenügend belegt. Verschiedene Studien zeigten, dass die Vernachlässigung der Anfangswahrscheinlichkeit selbst dann nicht verschwände, wenn die Versuchspersonen die natürlichen Häufigkeiten aufgrund eigener Beobachtungen feststellten.⁷³⁸ Auch GIGERENZERS eigene Resultate belegten keineswegs, dass die Vernachlässigung der Anfangswahrscheinlichkeit *verschwände*, sondern nur, dass sie sich bei der Darstellung der Information als Häufigkeiten *verringere*.
- 443 Die Kontroverse zwischen Vertretern der *heuristics and biases* Schule und den Vertretern der „evolutionären“ Sicht – worunter ich hier alle verstehe, die sich GIGERENZERS und COSMIDES/TOOBYs Schlussfolgerungen anschließen – tobt heftig, wie nicht anders zu erwarten ist, wenn zwei ausserordentlich erfolgsverwöhnte Akademiker von einem äusserst eloquenten und fundierten Kritiker angegriffen werden.⁷³⁹

⁷³¹ COSMIDES/TOOBY, FN 680, 16. COSMIDES/TOOBY entwickeln das evolutionspsychologische Argument für die relativen Häufigkeiten detaillierter als GIGERENZER, siehe COSMIDES/TOOBY, FN 680, 14 ff.

⁷³² GIGERENZER, FN 678, 149 ff.

⁷³³ DAVID M. EDDY, Probabilistic Reasoning in Clinical Medicine: Problems and Opportunities, in: KAHNEMAN/SLOVIC/TVERSKY (Hrsg.), 249-267; siehe auch W. CASSCELLS/A. SCHOENBERGER/T. GRAYBOYS, Interpretation by Physicians of Clinical Laboratory Results, New England Journal of Medicine 1978, 999-1000.

⁷³⁴ ULRICH HOFFRAGE/GERD GIGERENZER, Using Natural Frequencies to Improve Diagnostic Inferences, Academic Medicine 1998, 538-540.

⁷³⁵ COSMIDES/TOOBY, FN 680, 26.

⁷³⁶ GIGERENZER, FN 678, 149.

⁷³⁷ KAHNEMAN/TVERSKY, FN 179, 586.

⁷³⁸ KAHNEMAN/TVERSKY, FN 179, 584.

⁷³⁹ Siehe nur den Schlagabtausch zwischen GERD GIGERENZER/ULRICH HOFFRAGE, How to Improve Bayesian Reasoning Without Instructions: Frequency Formats, Psychological Review 1995, 684-704 mit CHARLES LEWIS/GIDEON KEREN, On the Difficulties Underlying Bayesian Reasoning: A Comment on Gigerenzer and Hoffrage, Psychological Review 1999, 411-416 und die Antwort von GERD

- 444 Als Aussenseiter ist es immer einigermaßen schwierig, zu entscheiden, wer bei einem solchen intellektuellen Schlagabtausch die besseren Argumente hat. Einerseits haben KAHNEMAN und TVERSKY sicher Recht, wenn sie darauf hinweisen, dass es nicht weiter erstaunlich ist, dass 92 % der Versuchspersonen die richtige Antwort geben, wenn man sie dazu zwingt, eine grafische Darstellung des Problems zu zeichnen.⁷⁴⁰ Wenn man den Leuten die richtige Lösung erklärt, verstehen sie sie in der Regel; deswegen kann man die intuitive, abweichende Antwort auch als Fehler bezeichnen. Auch gibt es Studien, die zeigen, dass *base rate neglect* auch bei nicht verbalen Problemen und *natural sampling* auftritt⁷⁴¹ und Versuchspersonen nicht zwischen relevanten und irrelevanten *base rates* unterscheiden können,⁷⁴² was dagegen spricht, dass Menschen evolutionär gewachsene Denkprozesse zur Verarbeitung natürlicher Häufigkeiten haben, die weitgehend fehlerfrei funktionieren.
- 445 Andererseits sind die Belege dafür zahlreich, dass die Vernachlässigung der Basisrate bei Darstellung in natürlichen Häufigkeiten weitgehend verschwindet, und die theoretische Erklärung leuchtet ein. Der Ansatz von GIGERENZER und COSMIDES/TOOBY wird sofort plausibel, wenn man sich die Lösung eines konkreten Problem in Form eines Häufigkeitsbaums und in Form des Bayes-Theorem vor Augen hält (dazu hinten, S. 167 ff.). Eine faire Zusammenfassung des derzeitigen Diskussionsstandes besteht daher darin, zu sagen, dass die Anfangswahrscheinlichkeit bei der Darstellung der Information in natürlichen Häufigkeiten sehr viel weniger oft ignoriert wird als bei der Darstellung der Information in Wahrscheinlichkeiten.⁷⁴³ Daraus ergeben sich konkrete Empfehlungen für die gerichtliche Praxis.

IV. Das Bayes-Theorem und richterliche Beweiswürdigung

A. US-amerikanische Lehre

- 446 Das Für und Wider der Anwendung des Bayes-Theorems auf die richterliche Beweiswürdigung wird in den USA seit 30 Jahren intensiv diskutiert.⁷⁴⁴ Einige Aspekte

GIGERENZER/ULRICH HOFFRAGE, Overcoming Difficulties in Bayesian Reasoning: A Reply to Lewis and Keren (1999) and Mellers and McGraw (1999), *Psychological Review* 1999, 425-430.

⁷⁴⁰ So das Experiment 4.1 von COSMIDES/TOOBY, FN 680, 34.

⁷⁴¹ ADAM S. GOODIE/EDMUND FANTINO, An Experimentally Derived Base-Rate Error in Humans, *Psychological Science* 1995, 101-106.

⁷⁴² GERNOT D. KLEITER/MARIANNE KREBS, Do Subjects Understand Base Rates?, *Organizational Behavior and Human Decision Processes* 1997, 25-61.

⁷⁴³ So lässt sich KOEHLER, FN 633, 50, grob zusammenfassen. Siehe auch ROBERT HAMM, Underweighting of Base-Rate Information Reflects Important Difficulties People Have With Probabilistic Inference, *Psychology* 94.5.03. base-rate.7.hamm (1994) (www.cogsci.ecs.soton.ac.uk/psychology/raw/1994.volume.5/psyc.94.5.03.base-rate.7.hamm; besucht am 14. Januar 2005).

⁷⁴⁴ Befürwortend IRVING J. GOOD, *Probability and the Weighing of Evidence*, London 1950; MICHAEL FINKELSTEIN/WILLIAM FAIRLEY, *A Bayesian Approach to Identification Evidence*, *Harvard Law Review* 1970, 489-517; MICHAEL FINKELSTEIN/WILLIAM FAIRLEY, A Comment on „Trial by Mathematics“, *Harvard Law Review* 1971, 1801-1809; RICHARD LEMPERT, *Modeling Relevance*, *Michigan Law Review* 1977, 1021-1057; DAVID KAYE, *The Laws of Probability and the Laws of the Land*, *University of Chicago Law Review* 1979, 34-56; SAKS/KIDD, FN 183; ablehnend LAWRENCE TRIBE, FN 184; L. JONATHAN COHEN, *The Probable and the Provable*, London 1977; LEA BRILMEYER/LEWIS KORNHAUSER, *Review: Quantitative* 148

dieser Debatte betreffen spezifisch amerikanische Probleme, wie die Frage, ob Geschworene überhaupt fähig sind, einer Erläuterung des Bayes-Theorems zu folgen, oder ob sie dadurch nicht mehr verwirrt als aufgeklärt werden. Andere Befürchtungen sind, dass Geschworene die Bedeutung explizit statistischer Informationen übergewichten und andere, „weiche“ Faktoren ausser Acht lassen.⁷⁴⁵ Ein guter Teil der Kontroverse dreht sich um die Frage, ob eine Verurteilung ausschliesslich aufgrund eines statistischen Beweismittels zulässig sei, oder ob es individualisierender Beweismittel bedürfe. Unter dem Titel *naked statistical evidence* werden hypothetische Fälle wie das „*gatecrasher paradox*“ und die Gefangenen-Verschwörung erörtert. Bei der letzteren verabreden sich 24 Gefangene, einen Wärter zu ermorden. Nach dem Mord befinden sich im abgeschlossenen Gefängnishof 25 Gefangene, die als Täter in Frage kommen. Jeder einzelne von ihnen ist daher mit einer Wahrscheinlichkeit von 96 % der Täter; einer aber mit Sicherheit unschuldig. Genügt dies, um alle von ihnen zu verurteilen?⁷⁴⁶ Beim „*gatecrasher paradox*“ besuchen 1'000 namentlich bekannte Personen ein Rodeo, zu dem erwiesenermassen nur 490 Eintrittsbillette verkauft wurden. Darf man, wenn keinerlei weitere Beweismittel vorliegen, alle 1'000 Personen zur Nachzahlung des Eintrittspreises verurteilen, weil die Wahrscheinlichkeit, dass jeder Einzelne die Veranstaltung ohne zu bezahlen besucht hat, 0,51 beträgt (und nach amerikanischen Zivilrecht „*preponderance of the evidence*“, was gemeinhin als eine Wahrscheinlichkeit von mehr als 0,5 verstanden wird, für eine Gutheissung der Klage genügt)?⁷⁴⁷

- 447 Während also ein Teil dieser Debatte getrost ignoriert werden kann, sind andere Aspekte der amerikanischen Diskussion auch für Europa relevant. Zu Recht wird das Problem der a priori Wahrscheinlichkeiten thematisiert, auf das die subjektive Wahrscheinlichkeitstheorie bekanntlich keine Antwort hat, und die Komplexität der Berechnungen herausgestrichen, die notwendig sind, wenn man mit der wahrscheinlichkeitstheoretischen Beweiswürdigung ernst macht. Auf diese beiden Aspekte wird nachgehend eingegangen, nachdem kurz der Standpunkt der schweizerischen (soweit vorhanden) und deutschen Lehre zur Anwendung explizit Bayes'scher Induktion in Gerichtsverfahren beleuchtet wurde. Abschliessend wird darauf hingewiesen, dass die subjektive Wahrscheinlichkeitstheorie, unabhängig von ihrer Anwendung in der gerichtlichen Praxis, die Dogmatik des Beweisrechts bereichern kann – anerkanntermassen nicht gerade ein Gebiet, in dem in den letzten Jahren grosse Fortschritte gemacht wurden.

Methods and Legal Decisions, University of Chicago Law Review 1978, 116-151; CHARLES NESSON, Reasonable Doubt and the Permissive Inference: The Value of Complexity, Harvard Law Review 1979, 1187-1215. Diese Aufzählung ist selbstverständlich nicht vollständig, was in Anbetracht der kaum zu überschauenden Menge an Publikationen auch nicht möglich wäre. Sie führt jedoch die meist zitierten „Klassiker“ der Debatte auf.

⁷⁴⁵ TRIBE, FN 184, 1361. Empirische Studien bestätigen diese Befürchtung nicht, siehe FAIGMAN/BAGLIONI, FN 621, 1.

⁷⁴⁶ Siehe NESSON, FN 744, 1192 f.

⁷⁴⁷ L. JONATHAN COHEN, Subjective Probability and the Paradox of the Gatecrasher, Arizona State Law Journal 1981, 627-633. Eine vertiefte Analyse von COHENS Paradoxen findet sich bei BOURMISTROV-JÜTTNER, FN 692, 172 ff.

B. Schweizer Lehre

- 448 In der Schweiz haben sich ersichtlich nur CHAMPOD/TARONI – beides keine Juristen, sondern forensische Wissenschaftler – zur Anwendung des Bayes-Theorem bei der richterlichen Beweiswürdigung geäußert:⁷⁴⁸

Das Bayes-Theorem findet im Strafprozess eine ideale Anwendung. Es erlaubt, auf verlässliche Weise die Häufigkeiten und Wahrscheinlichkeiten zu berücksichtigen und die flagranten Fehler, die auf die Intuition zurückzuführen sind, zu vermeiden.

- 449 Schweizer Juristen, die sich zur Beweiswürdigung äussern, begnügen sich damit, festzuhalten, dass die richterliche Beweiswürdigung frei sei, es keine schematischen Regeln gebe und die Beweiswürdigung das Ergebnis einer Gesamtschau sei, wobei sich der Beweiswert eines Beweismittels aus den konkreten Umständen des Einzelfalls ergebe.⁷⁴⁹ Stets wird jedoch verlangt, dass die Beweiswürdigung rational und nachvollziehbar zu begründen sei.⁷⁵⁰

C. Deutsche Lehre

- 450 Abgelehnt wird die Anwendung des Bayes-Theorems von Vertretern der skandinavischen Beweiswertmethode wie PER OLOF EKELOF. Für ihn hat die Anfangswahrscheinlichkeit „kein kognitives Verhältnis zu dem Beweisthema“, sie hat „auch keinen Beweiswert an sich und [kann] nicht in demselben Sinn angewandt werden wie nach dem Bayes-Theorem“.⁷⁵¹
- 451 Die meisten deutschsprachigen Autoren, die sich mit dem Thema auseinandergesetzt haben, befürworten die Anwendung des Bayes-Theorem, wenn auch meist nicht als starre Regel, gemäss der alle Beweise mathematisch zu gewichten und das Beweisergebnis gleichsam zu errechnen ist, sondern als Modell, um rational und intersubjektiv überprüfbar argumentieren zu können.⁷⁵² BENDER/NACK formulieren dies so:⁷⁵³

⁷⁴⁸ CHAMPOD/TARONI, FN 710, 213.

⁷⁴⁹ So z.B. ALFRED BÜHLER, Beweiswürdigung, in: CHRISTOPH LEUENBERGER, Der Beweis im Zivilprozess, Bern 2000, 71-92, 87. MICHAEL NONN beerdigt in seiner Dissertation aus dem Jahr 1996 in erster Linie die legale Beweistheorie, die von den meisten schon seit längerem für tot gehalten wird, MICHAEL NONN, Die Beweiswürdigung im Zivilprozess unter besonderer Berücksichtigung des Kantons Basel-Landschaft, Basel 1996.

⁷⁵⁰ BÜHLER, FN 749, 89; GULDENER, FN 722, 322; HABSCHIED, FN 724, 398.

⁷⁵¹ GRÄNS, FN 688, 149; PER OLOF EKELOF, Beweiswert, in: WOLFGANG GRUNSKY/ROLF STÜRMER/GERHARD WALTER/MANFRED WOLF (Hrsg.), Festschrift für Fritz Baur, Tübingen 1981, 343-363, 358: „Statistisch gesehen liegt hier eine Ursprungswahrscheinlichkeit von 90 % vor, dass die Kuh dem A gehört. Meiner Ansicht nach ist dieses Verhältnis nur ein negatives Hilfsfaktum von marginaler Bedeutung. Zwar haben wir die Erfahrung, dass es in einer solchen Situation wahrscheinlicher ist, dass die Kuh dem A gehört. Die Verteilung der Kühe auf ihre Eigentümer ist aber mit der Kuh im Garten nicht kausal verbunden. Für sich allein betrachtet hat diese Verteilung daher so wenig Beweiswert, dass sie im Verhältnis zum geforderten Beweismass ohne Bedeutung ist“. EKELOF scheint, wie die Versuchspersonen von KAHNEMAN/TVERSKY, der *base rate* nur dann Bedeutung zuzumessen, wenn sie mit dem Ereignis kausal verknüpft ist, siehe TVERSKY/KAHNEMAN, FN 606, 60.

⁷⁵² So bereits LEMPERT, FN 744, 1025, auf die Kritik von TRIBE reagierend.

⁷⁵³ BENDER/NACK, FN 605, Rz. 400.

Sinn und Zweck der Anwendung des Theorem von Bayes auf den Indizienbeweis ist nicht die Berechnung von Wahrscheinlichkeiten im Prozess; sondern seine Anwendung als Modell, um den Denkprozess beim Indizienbeweis zu beschreiben, um rationaler argumentieren zu können.

452 Für GRÄNS⁷⁵⁴ sind die

Formeln [...] als heuristische Hilfsmittel deutlich einer blossen Intuition vorzuziehen. [...] Zu behaupten, dass die Formeln irgendeine unberechtigte Kontrolle über die Beweiswürdigung nehmen würden, heisst zu behaupten, dass man nicht bereit ist, die Rationalität und die logische Haltbarkeit seiner Würdigung zu überprüfen.

453 NELL weist darauf hin, dass selbst die Kritiker der „Mathematisierung“ insbesondere des Strafprozesses in den USA die Gültigkeit des Bayes-Theorems als eines theoretischen Modells für eine folgerichtige Bildung von Wahrscheinlichkeitsurteilen nicht in Frage stellen.⁷⁵⁵

454 Am weitesten geht BOURMISTROV-JÜTTNER, die die Zuordnung von Wahrscheinlichkeitswerten zu intuitiven Urteilen fordert, weil sie zu einer vorsichtigeren Entscheidungspraxis führe,⁷⁵⁶ die Präzisierung des intuitiven Urteils und eine bessere intersubjektive Nachvollziehbarkeit der einzelnen Urteilsbildung erlaube sowie die rationale Lösung eines juristischen Falles ermögliche.⁷⁵⁷

D. Stellungnahme

455 Mit der Mehrheit der deutschen und amerikanischen Lehre würde ich so weit wie BOURMISTROV-JÜTTNER nicht gehen. Zu verlangen, dass Gerichte ihre Urteile gleichsam errechnen, ist nicht nur unrealistisch, es würde auch eine mathematische Scheingenauigkeit vortäuschen, die nicht existiert. Aber das Bayes-Theorem als *Urteilsheuristik* ist auf jeden Fall der ungestützten Intuition vorzuziehen, die, wie die oben dargestellte psychologische Forschung zeigt, oft in die Irre führt. Wenn man verlangt, dass die Beweiswürdigung „die Gesetze der Logik nicht verletzen [darf]“⁷⁵⁸ und der Richter zu „gewissenhafter Schlussfolgerung aufgrund des Ergebnisses des Beweisverfahrens“⁷⁵⁹ verpflichtet ist, kann man sich der Folge kaum entziehen, dass das Bayes-Theorem der normative Standard ist, an dem die Kohärenz der Beweiswürdigung gemessen werden muss. Ausschliesslich das Bayes-Theorem erlaubt es, eine bestehende Überzeugung und neue Erkenntnisse widerspruchsfrei zu integrieren.

⁷⁵⁴ GRÄNS, FN 688, 160.

⁷⁵⁵ NELL, FN 687, 53.

⁷⁵⁶ Unter Hinweis auf empirische Studien von RITA JAMES SIMON/LINDA MAHAN, Quantifying Burdens of Proof: A View from the Bench, the Jury and the Classroom, Law and Society Review 1971, 319-330.

⁷⁵⁷ BOURMISTROV-JÜTTNER, FN 692, 195.

⁷⁵⁸ HABSCHEID, FN 724, 398.

⁷⁵⁹ GULDENER, FN 722, 322.

E. Das Problem der a priori Wahrscheinlichkeiten

- 456 Wenn das Bayes-Theorem die widerspruchsfreie Kombination von Anfangswahrscheinlichkeit und Merkmalswahrscheinlichkeit erlaubt und wir daher, wenn wir beide kennen, auf sicherem Boden sind, richtet sich die Aufmerksamkeit naturgemäss darauf, wie man zu Anfangswahrscheinlichkeit und Merkmalswahrscheinlichkeit gelangt. Das Bayes-Theorem gibt darauf bekanntlich keine Antwort. Nicht ganz zu Unrecht ist bemerkt worden, dass eine Theorie der Beweiswürdigung, die nur die Kombination bekannter Wahrscheinlichkeiten erlaubt, aber keine Hinweise darauf gibt, wie man zu diesen Wahrscheinlichkeiten gelangt, unvollständig und für die Praxis wenig nützlich ist.

1. Die massgebliche Referenzklasse

- 457 COHEN argumentiert, dass die Versuchspersonen in KAHNEMAN/TVERSKYS Taxi-Problem die *base rate* von 85 % grünen Taxis zu Recht ignoriert haben:⁷⁶⁰

Warum nur soll man annehmen, dass Versuchspersonen, die gebeten werden, die unbedingte Wahrscheinlichkeit zu schätzen, dass das Taxi, *das den Unfall verursacht hat*, Blau gewesen ist, die stadtweite Verteilung von grünen und blauen Taxis beachten, die bestenfalls dann relevant ist, wenn die zu entscheidende Frage ist, ob der Zeuge, der irgendwo ein Taxi gesehen hat, *nicht notwendigerweise in einen Unfall verwickelt*, ein blaues Taxi gesehen hat?

- 458 Für COHEN ist die stadtweite Verteilung von grünen und blauen Taxis ein sehr schwaches Fundament, um die relevante Häufigkeit zu schätzen.⁷⁶¹ Aber welches ist die „relevante Häufigkeit“? Gibt es nur eine relevante Häufigkeit? Sicher ist „grüne und blaue Taxis, die in Unfälle verwickelt waren“ eine relevantere Häufigkeit als die stadtweite Verteilung grüner und blauer Taxis, weil sie auf mehr Informationen beruht und insbesondere die Neigung der verschiedenfarbigen Taxis, in Unfälle verwickelt zu sein, berücksichtigt. Interessanterweise wird die Basisrate der Taxis, die in Unfälle verwickelt sind, bei der Lösung des Taxi-Problems von der Mehrheit der Versuchspersonen beachtet, nicht aber die stadtweite Verteilung der Taxis an sich.⁷⁶² Es gibt selbstverständlich weitere Referenzklassen, die noch mehr relevante Informationen berücksichtigen, beispielsweise „Taxis, die nachts in Unfälle verwickelt waren“ oder „Taxis, die nachts in diesem Stadtteil in Unfälle verwickelt waren“, oder gar „Taxis, die nachts in dieser Strasse in Unfälle verwickelt waren“. Während diese Referenzklassen immer mehr Eigenschaften mit dem zu beurteilenden Ereignis teilen und insofern die Umstände des Einzelfalles besser berücksichtigen,

⁷⁶⁰ L. JONATHAN COHEN, Are There Any A Priori Constraints on the Study of Rationality, Behavioral and Brain Sciences 1981, 359-380, 365: “[W]hy on earth should it be supposed that subjects, asked to estimate the unconditional probability that the cab *involved in the accident* was blue, ought to take into account a prior distribution of colours that would at best be relevant only if the issue at stake was just about the colour of a cab that was said to have been seen somewhere, *not necessarily in an accident*, and was taken to be blue?”

⁷⁶¹ COHEN, FN 760, 365.

⁷⁶² AMOS TVERSKY/DANIEL KAHNEMAN, Evidential Impact of Base Rates, in: Kahneman/Slovic/Tversky (Hrsg.), 153-160, 157. Nach TVERSKY/KAHNEMAN wird die Basisrate der Unfälle beachtet, weil sie eine kausale Beziehung zum Einzelfall suggeriert: die Fahrer der Taxis, die häufiger in Unfälle verwickelt sind, fahren gefährlicher.

werden sie auch immer kleiner, weshalb die Verlässlichkeit der Häufigkeitsschätzung abnimmt.⁷⁶³

- 459 Die relevante Referenzklasse gibt es demnach nicht. Eine spezifischere Häufigkeit reduziert die Unsicherheit zweiten Grades (ob die relative Häufigkeit auch auf den Einzelfall zutrifft), weil sie die Menge der Informationen reduziert, die die Schätzung beeinflussen kann.⁷⁶⁴ Andererseits erhöht sich die Gefahr, dass die beobachtete Häufigkeit nicht dem Grenzwert der relativen Häufigkeit entspricht, umso mehr, je kleiner die Stichprobe ist, in der die Häufigkeit beobachtet wird. PAUL E. MEEHL schlägt daher als Faustregel vor, als beste Referenzklasse immer die kleinste Klasse zu berücksichtigen, der das zu beurteilende Ereignis zugeordnet werden kann und deren Anzahl Mitglieder gross genug ist, um verlässliche Schätzungen der relativen Häufigkeit zu erlauben.⁷⁶⁵
- 460 Was aber, wenn es keine Angaben darüber gibt, wie viele grüne im Verhältnis zu blauen Taxis in nächtliche Unfälle verwickelt waren? Darf man dann die bekannte, aber unspezifische Referenzklasse „alle in der Stadt zugelassenen Taxis“ einfach ignorieren? Wenn das Ziel ist, ein möglichst genaues, d.h. dem tatsächlich Vorgefallenen entsprechendes, Urteil zu treffen, lautet die Antwort nein. Denn wenn man die stadtweite Verteilung nicht beachtet, muss man – mangels anderer Indizien – davon ausgehen, dass grüne und blaue Taxis gleich häufig in Unfälle verwickelt sind und somit die Anfangswahrscheinlichkeit, dass es sich um ein grünes Taxi gehandelt hat, 0,5 beträgt. Die Häufigkeit von grünen und blauen Taxis in der ganzen Stadt ist aber ein Indiz dafür, dass dies nicht zutrifft. Wenn 85 % der Taxis in der Stadt blau sind, ist die beste Schätzung – ohne zusätzliche Informationen – dass auch 85 % der in nächtliche Unfälle verwickelten Taxis blau sind.⁷⁶⁶ Jede Abweichung von dieser Basisrate ist zu begründen – die Gesellschaft mit den blauen Taxis könnte beispielsweise mittels der Einsatzpläne ihrer Fahrer nachweisen, dass sie nachts kaum Taxis in Betrieb hat, etc.
- 461 Darüber, ob eine Referenzklasse die richtige ist, d.h. ob deren Randbedingungen auch auf den zu beurteilenden Fall zutreffen, lässt sich trefflich streiten. Wenn man sich aber bewusst ist, warum und worüber man streitet, kann man wenigstens die richtigen Fragen stellen und die konkreten Randbedingungen der Referenzklasse, die die Entscheidung begründen, werden transparent gemacht und können diskutiert werden.⁷⁶⁷

2. Geschätzte Anfangswahrscheinlichkeiten, oder „garbage in, garbage out“?

- 462 Bei den obigen Ausführungen zur massgeblichen Referenzklasse wurde vorausgesetzt, dass man die relativen Häufigkeiten in den verschiedenen Referenzklassen kennt. In der Praxis wird man sich aber häufig mit der Situation konfrontiert sehen, dass man die relati-

⁷⁶³ Und natürlich ist es auch möglich, dass es zu den spezifischeren Referenzklassen keine statistischen Daten gibt, siehe dazu den nächsten Abschnitt.

⁷⁶⁴ JONATHAN J. KOEHLER/DANIEL N. SHAVIRO, Veridical Verdicts: Increasing Verdict Accuracy Through the Use of Overtly Probabilistic Evidence and Methods, Cornell Law Review 1990, 247-279, 260.

⁷⁶⁵ MEEHL, FN 152, 22.

⁷⁶⁶ KOEHLER/SHAVIRO, FN 764, 261.

⁷⁶⁷ BENDER/NACK, FN 605, Rz. 393.

ven Häufigkeiten nicht kennt, weil keine Daten dazu erhoben wurden. Der Richter kann daher nicht vermeiden, die relativen Häufigkeiten zu schätzen, weil er – anders als der Wissenschaftler – mangels Daten nicht einfach auf ein Urteil verzichten kann. Er ist gezwungen, auch dann zu urteilen, wenn die empirische Datenbasis ungenügend ist.

- 463 Oft wird nun der Vorwurf erhoben, dass damit die ganze auf der Wahrscheinlichkeitstheorie beruhende Beweislehre wertlos ist, weil ihr Ausgangspunkt auf subjektiven, nicht empirisch erhärteten Schätzungen beruht. BENDER/NACK bemerken dazu Folgendes:⁷⁶⁸

Wenn man die Forderung aufstellt, dass der Richter die drei Fragen der Fragen-Trias [es handelt sich um die vorne, S. 134, wiedergegebenen Fragen] stellen – und beantworten! – muss, wird einem oft entgegengehalten, dass man über die Häufigkeitsverteilungen in der Regel keine Aussagen – und schon gar keine empirisch abgesicherten Aussagen – machen kann. Nur: Deswegen werden die richtigen Fragen nicht falsch. Denn allein das wäre die Alternative: Wenn man nichts über Häufigkeiten weiss und sie noch nicht einmal abschätzen kann, dann kann man logischerweise auch nicht entscheiden, ob ein Indiz belastend ist oder nicht. Daran, dass man oft sehr wenig über die Häufigkeitsverteilungen weiss, kann auch die Beweislehre nichts ändern. Es führt aber kein Weg daran vorbei, dass der Richter sich wenigstens auf Plausibilitätsniveau Vorstellungen über Häufigkeitsverteilungen machen muss, wenn er die Beweiskraft eines Indizes abschätzen will.

- 464 MUSIELAK sieht das ganz ähnlich:⁷⁶⁹

Mögen Schätzungen der Wahrscheinlichkeit aufgrund von Erfahrungssätzen des täglichen Lebens noch so unzulänglich erscheinen, insbesondere wenn man sie mit den beim forensischen Beweis die Ausnahme bildenden wissenschaftlich fundierten Wahrscheinlichkeitsberechnungen mit mathematischer Genauigkeit vergleicht, sie sind einer völlig unkontrollierten Ermessensentscheidung auf jeden Fall vorzuziehen.

- 465 Empirische Studien belegen, dass die Zerlegung eines Beweisthemas in einzelne Beweisfakten und die Abschätzung des abstrakten Beweiswertes (Likelihood-Quotienten) von einzelnen Indizien die Qualität der Entscheide selbst dann erhöht, wenn die Likelihood-Quotienten ohne empirische Grundlage geschätzt werden.
- 466 DAVID A. SCHUM und ANNE W. MARTIN baten ihre Versuchspersonen, die Zeugenaussagen in zwölf erfundenen Straffällen auf ihren Beweiswert hin zu beurteilen. In der gesamtheitlichen Bedingung (*zero task-decomposition*, ZTD) wurden die Versuchspersonen gebeten, nach der sorgfältigen Lektüre aller vorhandenen Zeugenaussagen den abstrakten Beweiswert aller Zeugenaussagen in ihrer Gesamtheit zu beurteilen, d.h. anzugeben, ob die Aussagen wahrscheinlicher sind, wenn der Angeklagte schuldig oder wenn er unschuldig ist. Eine Angabe von „*Guilt-10*“ bedeutet beispielsweise, dass die Zeugenaussagen zehn Mal wahrscheinlicher sind, wenn der Angeklagte schuldig ist, als wenn er unschuldig ist.
- 467 In der teilweise aufgeschlüsselten Bedingung (*partial task-decomposition*, PTD) wurden die Versuchspersonen gebeten, die gleichen Likelihood-Quotienten wie in der ZTD-Bedingung anzugeben, aber dieses Mal nicht für die Gesamtheit aller Aussagen, sondern für jede

⁷⁶⁸ BENDER/NACK, FN 605, Rz. 392.

⁷⁶⁹ HANS-JOACHIM MUSIELAK, Das Överviktsprincip: Zum Verhältnis von richterlicher Überzeugung und Wahrscheinlichkeit, in: ALEXANDER LÜDERITZ/JOCHEN SCHRÖDER, Festschrift für Gerhard Kegel, Frankfurt 1977, 451-471, 461.

Zeugenaussage separat. Das Bayes-Theorem erlaubt es, diese Beurteilungen der abstrakten Beweiskraft einzelner Zeugenaussagen zu einem Likelihood-Quotienten der Beweiskraft der gesamten Aussagen zu kombinieren.⁷⁷⁰

- 468 In der komplett aufgeschlüsselten Bedingung (*complete task-decomposition*, CTD) wurden die Versuchspersonen gebeten, nicht nur jede Zeugenaussage für sich, sondern einzelne Elemente jeder Zeugenaussage zu bewerten. Das gesamte Beweismaterial wurde also in einzelne Aussagen zerlegt, die dann je einzeln danach bewertet werden mussten, ob sie wahrscheinlicher sind, wenn der Angeklagte unschuldig ist oder wenn er schuldig ist. Beispielsweise wurden die Versuchspersonen gefragt, (i) wie wahrscheinlich eine Zeugenaussage ist, wenn der Angeklagte am Tatort *und schuldig* ist; (ii) wie wahrscheinlich die Aussage ist, wenn er am Tatort *und unschuldig* ist; u. s. w. Auch diese Likelihood-Quotienten lassen sich wieder mittels des Bayes-Theorem zu einem einzigen Likelihood-Quotienten für das gesamte Beweisergebnis kombinieren – der Leser geht recht in der Annahme, dass sowohl die Aufteilung der Zeugenaussagen in einzelne Elemente wie auch die Rekombination mittels Bayes-Theorem ausserordentlich komplex sind.⁷⁷¹
- 469 SCHUM/MARTIN haben die von ihnen erhobenen Daten auf verschiedene Art und Weise analysiert. Stark vereinfacht lassen sich ihre Ergebnisse wie folgt zusammenfassen: je feiner die zu fällenden Urteile über den Sachverhalt in einzelne Elemente aufgeteilt werden, desto eher waren sich die Versuchspersonen über die Beurteilung der Fälle einig.⁷⁷² D.h., je mehr den Versuchspersonen die Aufgabe der Aggregation und Integration der Daten abgenommen wurde und sie im Gegenzug dafür mehr Likelihood-Quotienten einzelner Aussageelemente abschätzen mussten, desto eher stimmte ihre (durch das Modell errechnete) Gesamtbeurteilung der Fälle überein. Wenn man davon ausgeht, dass grössere Übereinstimmung in der Beurteilung der Fälle ein Qualitätskriterium ist – d.h. wenn man annimmt, dass ein Fall eher „richtig“ entschieden wurde, wenn sich die meisten Urteiler über das Urteil einig sind – lässt dies den Schluss zu, dass Menschen weniger mit der Beurteilung einzelner Likelihood-Quotienten Mühe haben, sondern vielmehr mit der Integration grosser Mengen von teilweise widersprüchlichen Informationen. Diese Integration kann ihnen das Bayes-Theorem abnehmen. Über Beurteilung der Likelihood-Quotienten, die vom Bayes-Theorem als gegeben vorausgesetzt werden, scheinen sich Menschen eher einig zu werden als über die korrekte Integration der einzelnen Beweismittel zu einem Gesamtbild.
- 470 SCHUM/MARTIN können auch nachweisen, dass ihre Versuchspersonen bei einer Gesamtbeurteilung von teilweise widersprüchlichen Aussagen, Aussagen, die der ersten Aussage widersprachen, schlicht ignorieren oder gar als Bestätigung der ersten Aussage auffassen.⁷⁷³ Erst wenn die widersprüchliche Aussage in ihre einzelnen Elemente aufgeteilt

⁷⁷⁰ DAVID A. SCHUM/ANNE W. MARTIN, Formal and Empirical Research on Cascaded Inference, Law and Society Review 1982, 105-151; Nachdruck in: REID HASTIE (Hrsg.), Inside the Juror, Cambridge 1993, 136-174, 156. Die Seitenzahlen beziehen sich auf den Nachdruck.

⁷⁷¹ Die bei SCHUM/MARTIN, FN 770, 169 ff. abgedruckten Formeln geben einen Eindruck von der Komplexität des Modells; ebenso die Tatsache, dass die 20 Versuchspersonen insgesamt 16'000 Likelihood-Quotienten zu Protokoll gaben, SCHUM/MARTIN, 153.

⁷⁷² SCHUM/MARTIN, FN 770, 158 ff.

⁷⁷³ SCHUM/MARTIN, FN 770, 164.

und diese beurteilt werden, zeigt der abstrakte Beweiswert der zweiten, widersprüchlichen Aussage in die entgegengesetzte Richtung der ersten Aussage. SCHUM/MARTIN fanden also einen lokalen Primateneffekt (*primacy*), d.h. das zuerst gehörte Argument beeinflusste die Interpretation späterer Argumente (mehr dazu hinten, S. 188 ff.).

- 471 Andererseits wurden redundante Informationen in der ZTD-Bedingung doppelt gewichtet, wodurch ihnen ein unzulässig hohes Gewicht beigemessen wird,⁷⁷⁴ was eine entsprechende Befürchtung von LEMPERT bestätigt.⁷⁷⁵
- 472 Redundante Indizien sind solche, deren Informationsgehalt ganz oder teilweise in bereits berücksichtigten Indizien enthalten ist.⁷⁷⁶ Den Einfluss untereinander *unabhängiger* Indizien auf die Wahrscheinlichkeit der Hypothese kann man einfach berechnen, indem man die anfänglichen Chancen der Hypothese mit dem Likelihood-Quotienten des Indizes multipliziert und das Ergebnis als anfängliche Chance für die Multiplikation mit dem Likelihood-Quotienten des nächsten Indizes verwendet.⁷⁷⁷ Man wendet das Bayes-Theorem einfach mehrmals an und verwendet jeweils das Resultat der vorangehenden Anwendung als Anfangswahrscheinlichkeit für die nächste Anwendung.
- 473 Die Merkmalswahrscheinlichkeit zweier untereinander unabhängiger Indizien ergibt sich aus der Produktregel:

$$P(I_1 \text{ und } I_2) = P(I_1) \cdot P(I_2) \quad (18)$$

- 474 Die Anwendung der Produktregel ist aber unzulässig, wenn die beiden Indizien untereinander abhängig sind, d.h. wenn die Wahrscheinlichkeit, dass das zweite Indiz vorliegt, sich ändert, wenn das erste Indiz feststeht. Bei abhängigen Indizien ist die Merkmalswahrscheinlichkeit

$$P(I_1 \text{ und } I_2) = P(I_1) \cdot P(I_2 | I_1). \quad (19)$$

- 475 Wenn $P(I_2 | I_1) \approx 1$ verringert sich die Merkmalswahrscheinlichkeit bei Vorliegen des zweiten Indizes praktisch nicht mehr. Ein Beispiel von BENDER/NACK veranschaulicht diese Überlegung:⁷⁷⁸ angenommen, es steht fest, dass der Täter grosse Hände (I_1) und Füsse (I_2) hatte. Der Verdächtige hat ebenfalls grosse Füsse, und man weiss, dass derart grosse Füsse nur bei 10 % oder weniger der Gesamtbevölkerung vorkommen. Das Indiz „grosse Füsse“ tritt daher bei der Annahme, dass der Verdächtige der Täter ist, zehn Mal häufiger auf als bei der Annahme seiner Unschuld. Der Verdächtige hat, wie der Täter, auch grosse Hände, und man weiss, dass nur 5 % der Bevölkerung derart grosse Hände haben. Wären die beiden Indizien unabhängig, könnte man $0,1 \cdot 0,05 = 0,005$ (0,5 %) rechnen; mit anderen Worten ist die Wahrscheinlichkeit, dass die Indizienkombination bei

⁷⁷⁴ SCHUM/MARTIN, FN 770, 165.

⁷⁷⁵ LEMPERT, FN 744, 1050.

⁷⁷⁶ LEMPERT, FN 744, 1042.

⁷⁷⁷ BENDER/NACK, FN 605, Rz. 409.

⁷⁷⁸ BENDER/NACK, FN 605, Rz. 410.

Unschuld des Täters vorkommt, 200 Mal geringer als die Wahrscheinlichkeit, dass sie bei Schuld vorliegt.

- 476 Die Anwendung der Produktregel ist aber in diesem Fall offensichtlich unzulässig, denn die Indizien „grosse Füße“ und „grosse Hände“ sind nicht unabhängig; wer grosse Füße hat, hat in der Regel auch grosse Hände, und umgekehrt. Wenn nun 90 % der Menschen, die grosse Füße haben, auch grosse Hände haben, dann ist die Wahrscheinlichkeit ($I_2 | I_1$) = 0,9. Die Merkmalswahrscheinlichkeit von 0,1, die sich aus dem Indiz „grosse Füße“ ergibt, verringert sich daher nur auf $0,1 \cdot 0,9 = 0,09$ (von 10 % auf 9 %), wenn der Verdächtige auch grosse Hände hat.⁷⁷⁹
- 477 Abhängige Indizien sollten daher zu „Indizienfamilien“ zusammengefasst werden, deren Merkmalswahrscheinlichkeit gesamthaft beurteilt wird.⁷⁸⁰
- 478 Die Abhängigkeit zweier Indizien ist aber nicht immer so offensichtlich wie im obigen Beispiel mit den Händen und Füßen und wird bei der „Gesamtwürdigung“ der Beweise häufig übersehen, wie Ergebnisse von SCHUM/MARTIN belegen.⁷⁸¹ Schnell passiert es, dass man Indizien, die die Anfangswahrscheinlichkeit begründet und zur Verhaftung des Verdächtigen geführt haben, bei der Beweismittelwürdigung nochmals berücksichtigt.⁷⁸² Wenn die Polizei Falschgeld findet, dass sehr ähnlich ist wie dasjenige, das ein bekannter, vorbestrafter Fälscher jeweils herstellte, hat sie guten Grund, dem Mann einen Besuch abzustatten, denn es besteht ein begründeter Anfangsverdacht. Angenommen es finden sich keine weiteren Indizien für eine Täterschaft des Verdächtigen, aber der Staatsanwalt erhebt trotzdem Anklage. Aufgrund der Vorstrafe des Verdächtigen und des Falschgelds, das seine „Handschrift“ trägt, darf das Gericht vielleicht von einer Anfangswahrscheinlichkeit von 0,1 ausgehen, dass der Verdächtige schuldig ist. Es darf dann aber einen Schuldspruch nicht damit begründen, dass die Wahrscheinlichkeit, dass die Ähnlichkeit der Noten vorliegt, 100 Mal grösser ist, wenn der Angeklagte schuldig ist als wenn er unschuldig ist, woraus sich bei einer Anfangswahrscheinlichkeit von 0,1 eine Wahrscheinlichkeit der Schuld von 0,92 (92 %) ergibt, was möglicherweise für einen Schuldspruch genügen würde. Da die Ähnlichkeit der Noten bereits berücksichtigt wurde, um die Anfangswahrscheinlichkeit zu begründen, ist es unzulässig, sie nochmals zu berücksichtigen. Bei der Gesamtwürdigung der Beweismittel geschehen aber genau solche Fehler offenbar häufig.
- 479 Die Studien von SCHUM/MARTIN lassen daher den Schluss zu, dass die Qualität richterlicher Urteile auch dann verbessert werden kann, wenn die Likelihood-Quotienten für die einzelnen Beweiselemente ohne empirische Basis geschätzt werden, aber die Integration der Information der Regel Bayes' folgt.

⁷⁷⁹ Weitere Beispiele bei LEMPert, FN 744, 1043.

⁷⁸⁰ LEMPert, FN 744, 1044; BENDER/NACK, FN 605, Rz. 410.

⁷⁸¹ Daher ist es fraglich, ob die schweizerische Lehre mit ihrer Betonung der „Gesamtwürdigung“ der Beweise (BÜHLER, FN 749, 87) viel dazu beiträgt, die Qualität der Beweismittelwürdigung zu verbessern.

⁷⁸² LEMPert, FN 744, 105; BENDER/NACK, FN 605, Rz. 420.

F. Dogmatische Erkenntnisse

- 480 Die wahrscheinlichkeitstheoretische Analyse der Beweiswürdigung hat auch unabhängig davon, dass sie Fehler zu vermeiden hilft, Vorteile für die dogmatische Durchdringung des Beweisrechts. Welches Beweismass gefordert ist, was „Unschuldsvermutung“ eigentlich bedeutet und wann ein Beweismittel „erheblich“ im Sinne des Gesetzes (§ 133 ZPO-ZH u.a.) ist, lässt sich klarer analysieren, wenn man sich der Strukturen der Wahrscheinlichkeitstheorie bedient.
- 481 Was bedeutet beispielsweise die in der schweizerischen Literatur herrschende Umschreibung des vollen Beweismasses, dass ein Sachverhalt „mit an Sicherheit grenzender Wahrscheinlichkeit“⁷⁸³ erstellt sein muss, damit der Richter von seiner Richtigkeit überzeugt sein darf? Wird die erforderliche Gewissheit nicht numerisch in einer subjektiven Wahrscheinlichkeit ausgedrückt, ist es möglich, dass zwei Personen zwar meinen, vom gleichen zu reden, aber eigentlich ein ganz anderes Verständnis vom verlangten Beweismass haben.⁷⁸⁴ Dass diese Befürchtung nicht nur theoretischer Natur ist, zeigen zahlreiche amerikanische Studien, in denen Richter, Studenten und Normalbürger den Begriffen „*preponderance of the evidence*“ und „*beyond reasonable doubt*“ sehr unterschiedliche Bedeutungen zumessen (siehe Tabelle 11, siehe GRÄNS, FN 688, 187 f. für weitere Studien auch aus Skandinavien).

Direkte Abfrage des numerischen Wahrscheinlichkeitswerts

Study	Subjects	Preponderance of evidence	Beyond reasonable doubt
Simon (1969)	Judges	.61	.89
Simon & Mahan (1971)	Citizens	.77	.79
	Students	.76	.89
Davis (1976)	Students		
	Stringent criterion instruction	-	.87
	Lax criterion instruction	-	.82
	Undefined criterion instruction	-	.82
US v. Fatico (1978, Weinstein memorandum)	Judges	.51	.86
Kassin & Wrightsman (1979)	Students	-	.87
McCauliff (1982) (also "Clear and convincing". ⁷⁵)	Judges	.55	.90
Hastie, Penrod, & Pennington (1983)	Citizens	-	.92
Cowan, Thompson, & Ellsworth (1984)	Citizens	-	.86
MacCoun (1984)	Students		
	Continuous scale	.68	.67
	Interval (0, 1, ..., 10) scale	.68	.69

⁷⁸³ VOGEL/SPÜHLER, FN 715, 256.

⁷⁸⁴ GRÄNS, FN 688, 164.

Odds ratio scale		.49	.51
Dane (1985)	Students	-	.66
Kerr, Bull, MacCoun, & Rathbum (1985)	Students	-	.79
Kagehiro (1986)	Students	.82	.91
Martin & Schum (1987) Odds ratio scale	Students		.91
MacCoun & Kerr (1988) (reported in MacCoun, 1987)	Students	-	.87
	Citizens	-	.83
	Students	.69	.81
Kagehiro (1990)	Students	-	.81-.91

Tabelle 11: mathematische Bedeutung unterschiedlicher Beweismasse⁷⁸⁵

482 Noch interessanter sind Studien, in denen die entsprechenden Werte indirekt erhoben wurden. RITA JAMES SIMON hat zwei unabhängige Gruppen von Studenten gebeten, nach der Anhörung der Plädoyers in einem Straffall ihr Urteil über die Schuld des Angeklagten abzugeben. Das verlangte Beweismass war „*guilty beyond reasonable doubt*“; die Versuchspersonen in der einen Gruppe wurden gebeten, anzugeben, ob der Angeklagte schuldig im Sinne der Anklage sei oder nicht. Die Versuchspersonen der anderen Gruppe wurden gebeten, die Wahrscheinlichkeit anzugeben, dass der Angeklagte sich so verhalten hat, wie die Anklage ihm vorwirft. Die Wahrscheinlichkeitswertungen wurden der Grösse nach geordnet und die dem Kriterium „*guilty beyond reasonable doubt*“ entsprechende Wahrscheinlichkeit ermittelt, indem die Wahrscheinlichkeitswertungen um die Anzahl der Schuldsprüche in der anderen Gruppe heruntergezählt wurden. Der so erhaltene Wert lag bei 0,7, also erheblich tiefer als bei der direkten Abfrage.⁷⁸⁶ Die Resultate von SIMON konnten von MACCOUN (0,56), DANE (0,72) und zahlreichen weiteren (siehe Tabelle 12) repliziert werden.⁷⁸⁷

Indirekte Abfrage des numerischen Wahrscheinlichkeitswerts

Study	Subjects	Preponderance of evidence	Beyond reasonable doubt
Nagel, Lamm, & Neef (1981)	Students	-	.55
Nagel (1979)			
No instruction	Students	-	.60
Beyond reasonable doubt instruction		-	.68
.90 probability instruction		-	.80
"10-to-1 Blackstone error ratio" instruction		-	.90
MacCoun (1984)	Students	.52	.56
MacCoun (1994) ('Blackstone ratio')	Students	.48	.51

⁷⁸⁵ Aus REID HASTIE, Algebraic Models of Juror Decision Processes, in: HASTIE, FN 770, 84-115, 102.

⁷⁸⁶ RITA JAMES SIMON, „Beyond reasonable doubt“ – An Experimental Attempt at Quantification, Journal of Applied Behavioral Science 1970, 203-209.

⁷⁸⁷ Nachweise bei HASTIE, FN 785, 103.

Besonderer Teil

Thompson, Cowan, Ellsworth & Harrington (1984)	Citizens		
Prodeath penalty citizens	-		.50
Antideath penalty citizens	-		.58
Dane (1985)	Students	-	.52
MacCoun & Kerr (reported in MacCoun, 1987)	Citizens	-	.55
	Students	-	.52
	Students	.52	.54
MacCoun & Tyler (reported in MacCoun, 1987)	Citizens	-	.54
	Students	-	.51

Tabelle 12: Numerische Wahrscheinlichkeitswerte bei indirekter Abfrage (*parallel counting*)⁷⁸⁸

- 483 Diese Ergebnisse lassen nur den Schluss zu, dass Formulierungen wie „mit an Sicherheit grenzender Wahrscheinlichkeit“ von verschiedenen Personen gänzlich unterschiedlich verstanden werden und schon deshalb kein geeignetes Kriterium zur Festlegung des Beweismasses sind. Die Resultate der Studien mit indirekten Abfragemethoden bestätigen zudem den subjektiven Eindruck, dass es für eine Verurteilung häufig weniger braucht, als man gemeinhin annimmt – und insbesondere auch weniger, als die urteilenden Personen auf direkte Nachfrage hin zuzugeben bereit wären.⁷⁸⁹
- 484 Was bedeutet es, dass jede Person, die einer Straftat angeklagt ist, bis zum gesetzlichen Beweis ihrer Schuld als unschuldig gilt (Art. 6 Ziff. 2 EMRK)? Wird damit nur die Beweislastverteilung geregelt?⁷⁹⁰ Oder wird damit auch etwas über die zulässige Anfangswahrscheinlichkeit gesagt, die der Richter seiner Beweiswürdigung zu Grunde legen darf? Darf die Richterin mit anderen Worten von einer Anfangswahrscheinlichkeit der Schuld des Angeklagten von 0,9 ausgehen, wenn sie weiss – wie es empirisch zweifellos zutrifft – dass 90 % aller Anklagen in einem Schuldspruch enden?⁷⁹¹ Sagt die Unschuldsvermutung möglicherweise auch etwas über die Anfangswahrscheinlichkeit, die ein Gericht der Schuld des Angeklagten beimessen darf? Darf man den Kreis der möglichen Täter durch das Indiz des Anfangsverdachts eingrenzen, d.h. darf man davon ausgehen, dass der Angeklagte durch das Los aus dem Kreis aller Personen, auf die der Anfangsverdacht zutreffen kann, bestimmt wurde?⁷⁹² Oder muss das Gericht vor Beginn der Beweiswürdigung von einer Chance der Schuld von 1 zur Gesamtbevölkerung ausgehen,

⁷⁸⁸ Aus HASTIE, FN 785, 105.

⁷⁸⁹ Siehe zum Thema auch die finisch-schwedische Studie von MINNA HATAKKA/HANNU TAPANI KLAMI, Beweismass und Irrtumsrisiko, Helsinki 1990.

⁷⁹⁰ NIKLAUS SCHMID, Strafprozessrecht, 2. Aufl., Zürich 1993, Rz. 278 ff.

⁷⁹¹ WILLEM A. WAGENAAR, The Subjective Probability of Guilt, in: AYTON/WRIGHT, FN 619, 529-547, 531 (WAGENAAR stellt die Frage, er bejaht sie nicht).

⁷⁹² So BENDER/NACK, FN 605, Rz. 420, die allerdings betonen, dass man dann das Indiz, das den Anfangsverdacht begründete und zur Eingrenzung des möglichen Täterkreises (und damit zur erhöhten Anfangswahrscheinlichkeit) führte, bei der Beweiswürdigung nicht nochmals zählen darf.

um den Angeklagten nicht für schuldiger als jeden anderen zu halten; m. a. W. muss das Gericht den Angeklagten so behandeln, als wäre er zufällig per Los zum Angeklagten bestimmt worden (wobei die gesamte Bevölkerung an der Verlosung teilgenommen hat)?⁷⁹³ Wenn ausschliesslich die Richtigkeit des Urteils das massgebliche Kriterium ist, weckt diese Betrachtungsweise allerdings Bedenken – es werden möglicherweise zu viele Angeklagte freigesprochen. Oder sind diese Überlegungen ohnehin verfehlt, weil man aus dem Sollen, das das Recht vorschreibt, kein Sein der tatsächlichen Anfangswahrscheinlichkeit ableiten darf?⁷⁹⁴

- 485 Diese Fragen können hier nicht vertieft werden. Sie zeigen aber, welchen Erkenntnisgewinn die Anwendung der Wahrscheinlichkeitstheorie auf die Beweiswürdigung bieten kann; ganz abgesehen davon, dass sie die Urteilsqualität verbessern kann.

V. Eigene Studie zur Repräsentativitätsheuristik

A. Frage zur Repräsentativitätsheuristik

- 486 Den Richtern der Zivil- und Straferichte der Kantone Aargau, St. Gallen und Zürich wurde folgender Sachverhalt zur Beurteilung vorgelegt (2003):

Nachdem auch ohne Anlass Alkoholtests durchgeführt werden dürfen, lässt die Kantonspolizei Zürich alle Autofahrer, die eine Kontrollstelle am Stadtrand von Zürich passieren, ins Teströhrchen blasen. Erfahrungsgemäss ist einer von hundert getesteten Autofahrern betrunken.

Dieser Alcotest zeigt nicht ganz zuverlässig an, ob ein Autofahrer mehr als die zulässige Blutalkoholkonzentration aufweist. Es kommt vor, dass ein alkoholisierte Autofahrer nicht erwischt wird, weil sich das Röhrchen (leider) nicht verfärbt. Bedauerlicherweise verfärbt sich das Röhrchen aber auch gelegentlich bei nüchternen Autofahrern.

- Die Verfärbung tritt bei 95 % aller Blaufahrer (zu Recht) ein.
- Die Verfärbung tritt aber auch (bedauerlicherweise) bei 5 % aller nüchternen Fahrer ein.

Angenommen, einer von den vielen kontrollierten Autofahrern bläst ins Röhrchen, und dieses verfärbt sich.

Wie hoch ist die Wahrscheinlichkeit, dass die Polizei einen Blaufahrer erwischt hat?

- 487 Der Sachverhalt entspricht einem Beispiel von BENDER/NACK,⁷⁹⁵ ergänzt um die Anfangswahrscheinlichkeit⁷⁹⁶ („Erfahrungsgemäss ist einer von hundert getesteten Autofahrern betrunken“). Die Richter hatten vier Antwortmöglichkeiten (0-25 %, 26-50 %, 51-75 %, 76-100 %) und wurden gebeten, einen Prozentbereich zu wählen.

- 488 Die bedingte Wahrscheinlichkeit, dass der Test verfärbt, wenn der Autofahrer betrunken ist, beträgt 0,95. Die Anfangswahrscheinlichkeit, dass der Autofahrer betrunken ist, beträgt

⁷⁹³ So RICHARD LEMPert, The New Evidence Scholarship: Analyzing the Process of Proof, Boston University Law Review 1986, 439-477, 464.

⁷⁹⁴ LEMPert, FN 793, 463.

⁷⁹⁵ BENDER/NACK, FN 605, Rz. 435.

⁷⁹⁶ Die tatsächliche Anfangswahrscheinlichkeit ist leider höher und liegt bei 3-5%, siehe Berner Zeitung vom 24. November 2003, S. 24b und Neue Luzerner Zeitung vom 18. November 2003, S. 7.

Besonderer Teil

0,01. Die bedingte Wahrscheinlichkeit, dass der Test positiv anzeigt, wenn der Autofahrer nüchtern ist, beträgt 0,05. Die bedingte Wahrscheinlichkeit, dass der Autofahrer betrunken ist, wenn der Test verfärbt beträgt daher $(0,95 \cdot 0,01) / ((0,95 \cdot 0,01) + (0,05 \cdot 0,99)) = 0,16$ oder 16 %.

- 489 Die Hypothese ist, dass die Richter sehr viel häufiger die Antwort „76-100 %“ als die korrekte Antwort „0-25 %“ geben, da erstere die Antwort ist, die bei Verwechslung der bedingten Wahrscheinlichkeiten $P(\text{Test pos.} \mid \text{betrunken}) = 0,95$ und $P(\text{betrunken} \mid \text{Test pos.}) = 0,16$ zu erwarten ist.

B. Resultate

- 490 173 der 181 Richter, die den Fragebogen zurückschickten, beantworteten die Frage nach der Belastungswahrscheinlichkeit. 8,8 % (16) beantworteten die Frage richtig, indem sie 0-25 % wählten, 10,5 % (19) kreuzten 26-50 % an, 7,7% (14) wählten 51-75 % und 68,5 % (124) entschieden sich 76-100 %. Viele Richter gaben daher die Antwort, die bei der Verwechslung der bedingten Wahrscheinlichkeiten zu erwarten ist.
- 491 Die Richter schnitten damit besser ab als die Ärzte in der Studie von DAVID M. EDDY, die zu 95 % den Fehler begingen, die bedingten Wahrscheinlichkeiten zu verwechseln.⁷⁹⁷ Verglichen mit den amerikanischen Richtern in der Studie von GUTHRIE/RACHLINSKI/WISTRICH schnitten die schweizerischen Richter etwas schlechter ab. 40 % der amerikanischen Richter beantworteten die Frage richtig, ebenfalls 40 % gaben eine Antwort, die auf die Verwechslung der bedingten Wahrscheinlichkeiten hindeutet.⁷⁹⁸ Drei der Richter, die die Frage richtig beantworteten, zeichneten einen Häufigkeitsbaum nach Abbildung 10 und bewiesen damit, dass sie das Problem erkannt und verstanden hatten.

C. Diskussion

- 492 Die Resultate zeigen, dass eine grosse Mehrheit der befragten Richterinnen die (geringe) Anfangswahrscheinlichkeit nicht genügend gewichteten und – sei es als Folge davon oder aufgrund der semantischen Ähnlichkeit der inversen Wahrscheinlichkeiten – die bedingten Wahrscheinlichkeiten $P(\text{Test pos.} \mid \text{betrunken})$ und $P(\text{betrunken} \mid \text{Test pos.})$ verwechselten.
- 493 Nicht völlig ausgeschlossen werden kann, dass einige der Richter unbewusst eine höhere Anfangswahrscheinlichkeit annahmen. Obwohl die Aufgabenstellung klar stellt, dass die Polizei die Alkoholproben ohne Anlass durchführte, haben die Richter ihrer Entscheidung möglicherweise das 2003 geltende Recht zugrunde gelegt, das Alkoholproben nur bei Anzeichen von Angetrunkenheit zulies.⁷⁹⁹ „Anzeichen von Angetrunkenheit“ kann geradezu als Umschreibung für erhöhte Anfangswahrscheinlichkeit verstanden werden. Bereits

⁷⁹⁷ Zumindest legte ihre Antwort nahe, dass sie diesen Fehler begingen, EDDY, FN 733, 249 ff.; siehe auch CASSCELLS/ SCHOENBERGER/GRAYBOYS, FN 733, 999.

⁷⁹⁸ GUTHRIE/RACHLINSKI/WISTRICH, FN 20, 809.

⁷⁹⁹ Art. 55 Abs. 2 aSVG. Mit den am 1. Januar 2005 in Kraft getretenen Änderungen des Strassenverkehrsgesetzes wurde die anlassfreie Atemalkoholkontrolle eingeführt.

bei einer Anfangswahrscheinlichkeit von 20 % erhöht sich die Belastungswahrscheinlichkeit bei positivem Atemalkoholtest auf über 76 %.⁸⁰⁰

- 494 Das etwas bessere Abschneiden der amerikanischen Richter kann möglicherweise dadurch erklärt werden, dass an amerikanischen Rechtsfakultäten schon seit den späten siebziger, frühen achtziger Jahren die Anwendung des Bayes-Theorems in der Beweiswürdigung gelehrt wird. Die Generation der Richter, die befragt wurde, war sich daher der Bedeutung der Anfangswahrscheinlichkeit möglicherweise bewusster als der durchschnittliche schweizerische Richter, der im Rahmen seiner Ausbildung kaum mit dem Bayes-Theorem konfrontiert wird.
- 495 An der Repräsentation der Basisrate in natürlichen Häufigkeiten oder Wahrscheinlichkeiten – die bekanntlich einen Einfluss auf die korrekte Integration der Anfangswahrscheinlichkeit haben kann – kann es auf jeden Fall nicht gelegen haben. Es wurde darauf geachtet, dass die Repräsentation der Information derjenigen bei GUTHRIE/RACHLINSKI/WISTRICH entsprach, nämlich natürliche Häufigkeit bei der *base rate* („einer von hundert“) und Prozentzahlen bei den bedingten Wahrscheinlichkeiten $P(\text{Test pos.} \mid \text{betrunken})$ und $P(\text{betrunken} \mid \text{Test pos.})$.⁸⁰¹

VI. Wie kann man Fehler des repräsentativen Denkens vermeiden?

A. Gutachter sollten nie die Belastungswahrscheinlichkeit angeben

- 496 Gerichtsgutachter in der Schweiz, die sich zur Wahrscheinlichkeit äussern, dass eine bestimmte Spur – sei es eine Faser, Glas oder DNA-Spur – dem Angeschuldigten zugeordnet werden kann, äussern sich oft zur Belastungswahrscheinlichkeit („die Wahrscheinlichkeit der Vaterschaft nach ESSEN-MÖLLER ist sehr viel grösser als 99,99 %. Die Bewertung nach HUMMEL daher: Vaterschaft praktisch erwiesen“).⁸⁰²
- 497 Aussagen zur Belastungswahrscheinlichkeit setzen aber wie gesagt Annahmen über die Anfangswahrscheinlichkeit voraus. Der Gutachter kann diese Annahmen in der Regel nicht treffen, weil er nicht den gesamten Sachverhalt kennt.⁸⁰³ Gutachter rechnen daher häufig mit einer „neutralen“ Anfangswahrscheinlichkeit von 50 %;⁸⁰⁴ diese Anfangswahrscheinlichkeit setzt auch die Formel nach ESSEN-MÖLLER voraus.⁸⁰⁵
- 498 Die „neutrale“ Anfangswahrscheinlichkeit von 50 % wirkt sich aber häufig stark zu Lasten des Angeschuldigten aus. Bei einem Abstammungsgutachten mag sie, in Anbetracht der Ausschlussleistung moderner DNA-Gutachten (insbesondere, wenn die Proben, wie bei Vaterschaftsgutachten, gut sind), angemessen sein. Die Mutter des Kindes wird während

⁸⁰⁰ Nämlich auf 82,6%.

⁸⁰¹ GUTHRIE/RACHLINSKI/WISTRICH, FN 20, 808, verwendeten einen anderen Sachverhalt und andere Zahlen, aber die gleiche Repräsentation.

⁸⁰² TARONI/AITKEN, FN 662, 298; MANGIN/TARONI/BÄR, FN 645, 440.

⁸⁰³ BÄR, FN 665, 46.

⁸⁰⁴ BÄR, FN 665, 46; TARONI/AITKEN, FN 662, 301, mit Nachweisen.

⁸⁰⁵ TARONI/AITKEN, FN 662, 301.

der möglichen Empfängniszeit, die relativ klar eingegrenzt werden kann, nicht mit einer beliebigen Vielzahl von Männern geschlafen haben. Die Anfangswahrscheinlichkeit ist daher bereits relativ gross, wenn man weiss, dass der mögliche Vater einer der Männer ist, die zum Empfängniszeitpunkt Geschlechtsverkehr mit der Mutter hatten.

- 499 Anders liegt der Fall aber bei Spuren, die am Tatort eines Verbrechens gesichert werden. Der Kreis der möglichen Täter ist dort erst einmal sehr gross, so dass man nicht einfach von einer Wahrscheinlichkeit von 50 %, dass der Verdächtige der Täter ist, ausgehen darf.⁸⁰⁶
- 500 Der Gutachter sollte daher richtigerweise nur die Merkmalswahrscheinlichkeit angeben, d.h. die Häufigkeit, mit der ein bestimmtes Merkmal in der Gesamtbevölkerung vorkommt.⁸⁰⁷
- 501 Wenn der Gutachter in solchen Fällen die Belastungswahrscheinlichkeit – BÄR spricht von „Identitätswahrscheinlichkeit“⁸⁰⁸ – angibt, wird der Richter auf den Holzweg geführt. Die ohnehin schon bestehende Gefahr, dass die Anfangswahrscheinlichkeit vernachlässigt wird, steigert sich praktisch zur Gewissheit, da der Richter kaum die Berechnungen des Gutachters nachprüfen und feststellen wird, dass diese von einer Anfangswahrscheinlichkeit von 50 % ausgegangen ist.
- 502 BÄR rechtfertigt die Annahme einer Anfangswahrscheinlichkeit von 50 % mit der folgenden Überlegung:⁸⁰⁹

Bei diesem Ansatz [Berechnung nach ESSEN-MÖLLER] geht – stillschweigend – ein neutrales a-priori (0,5) mit ein, d.h. man geht von der fiktiven Annahme aus, dass in jedem zweiten (!) Fall tatsächlich Identität vorläge. In Wirklichkeit überwiegen aber die Fälle mit Identität diejenigen mit Nichtidentität. Mit Bestimmtheit gilt dies bei Berücksichtigung der extrem hohen Ausschlussleistung der DNA-Analyse, da die Nichtverursacher der Spur ja durch Ausschluss „abgereichert“ werden.

- 503 Diese Argumentation ist nicht schlüssig, und BÄR ist zwischenzeitlich auch von ihr abgerückt.⁸¹⁰ BÄR hält selber fest, dass der Gutachter die tatsächliche Anfangswahrscheinlichkeit nicht kennen kann.⁸¹¹ Der Gutachter kann daher auch nicht wissen, dass die Fälle mit Identität überwiegen, die Anfangswahrscheinlichkeit also über 0,5 liegt. Die Argumentation von BÄR, dass sich die hohe Anfangswahrscheinlichkeit daraus ergebe, dass ein Grossteil der Bevölkerung durch die DNA-Analyse als Spurenverursacher ausgeschlossen werden kann („extrem hohe Ausschlussleistung der DNA-Analyse“), beruht auf einem Versehen. Das Argument von BÄR berücksichtigt die Merkmalswahrscheinlichkeit zwei Mal; einerseits, um die Anfangswahrscheinlichkeit festzulegen, und dann nochmals, um den – durch die Berücksichtigung der Merkmalswahrscheinlichkeit bereits stark verkleinerten Kreis der Verdächtigen – weiter einzugrenzen.

⁸⁰⁶ TARONI/AITKEN, FN 662, 300.

⁸⁰⁷ TARONI/AITKEN, FN 662, 302 f.

⁸⁰⁸ BÄR, FN 665, 45.

⁸⁰⁹ BÄR, FN 665, 46. Das Ausrufzeichen entspricht dem Originaltext.

⁸¹⁰ MANGIN/TARONI/BÄR, FN 645, 339 ff.

⁸¹¹ BÄR, FN 665, 46.

504 Ein Beispiel mag veranschaulichen, was passiert. Angenommen, als Täter kommen alle Männer in der Schweiz zwischen 20 und 64 in Frage, also rund 2,3 Millionen Menschen.⁸¹² Am Tatort wird ein Haar gefunden, das nicht dem Opfer gehört. Eine DNA-Analyse des Haares ergibt einen „genetischen Fingerabdruck“, der nur bei 0,1 Promille der männlichen Bevölkerung vorkommt. Die DNA eines zufällig aus der Gesamtheit der Männer in der Schweiz ausgewählten Mannes stimmt also mit einer Wahrscheinlichkeit von 0,0001 mit der am Tatort gefundenen DNA-Spur überein.

505 Wie gross ist die Wahrscheinlichkeit, dass ein x-beliebiger Mann der Täter ist? Die Chance beträgt rund 1:2,3 Millionen. Wie gross ist die Chance, dass ein Mann der Täter ist, wenn sein genetischer Fingerabdruck mit dem am Tatort gefundenen übereinstimmt? Nach dem Bayes-Theorem in seiner Chancen-Form

$$\frac{P(\text{Täter} \mid \text{DNA matches})}{P(\neg \text{Täter} \mid \text{DNA matches})} = \frac{P(\text{Täter})}{P(\neg \text{Täter})} \cdot \frac{P(\text{DNA matches} \mid \text{Täter})}{P(\text{DNA matches} \mid \neg \text{Täter})} \quad (20)$$

numerisch $(1/2'300'000) / (2'300'000/2'300'001) \times 1/0,0001 \approx 1:230$.

506 Die Wahrscheinlichkeit, dass ein zufällig ausgewählter Mann, *dessen DNA mit der Tatortspur übereinstimmt*, der Spurenger ist, beträgt also rund 0,004.

507 Mit anderen Worten hat sich die Chance von 1:2,3 Millionen verringert auf 1:230. Dies ist die – in der Tat beachtliche – „Ausschlussleistung“ des DNA-Fingerabdrucks: 99,99 % der männlichen Bevölkerung scheiden als Täter aus, weil ihre Allele keine identische DNA-Sequenz aufweisen. BÄR will nun diesen durch die Ausschlussleistung „abgereicherten“ Pool als Anfangswahrscheinlichkeit festlegen und die Übereinstimmung des DNA-Fingerabdrucks erneut berücksichtigen, um den Pool weiter zu verringern. Im Beispiel käme man damit auf eine Chance von 10'000:230, dass es sich bei dem zufällig ausgewählten Mann, dessen DNA-Fingerabdruck mit demjenigen der Tatortspur übereinstimmt, um den Täter handelt (was einer Belastungswahrscheinlichkeit von rund 98% entspricht und möglicherweise zur Verurteilung eines Mannes führt, der in Wirklichkeit nur mit einer Wahrscheinlichkeit von weniger als einem halben Prozent der Täter ist). BÄR berücksichtigt das den Anfangsverdacht begründende Indiz zwei Mal, wovon BENDER/NACK zu Recht warnen.⁸¹³

508 BÄR hat in einem 1995 publizierten Aufsatz noch bedauert, dass sich die Angabe von Identitätswahrscheinlichkeiten international in der forensischen Spurenkunde noch nicht durchgesetzt hat.⁸¹⁴ Inzwischen vertritt auch er die Meinung, dass die Angabe der Belastungswahrscheinlichkeit Gerichte leicht in die Irre führen kann.⁸¹⁵

⁸¹² Die Zahl stimmt ungefähr, siehe Bundesamt für Statistik, Statistisches Jahrbuch der Schweiz 2004, 34. Fragen kann man sich, ob man das Universum der möglichen Täter auf die ständige Wohnbevölkerung der Schweiz eingegrenzen darf, denn grundsätzlich kommt auch eine sich auf der Durchreise befindliche Person als Täter in Frage.

⁸¹³ BENDER/NACK, FN 605, Rz. 420.

⁸¹⁴ BÄR, FN 665, 45.

⁸¹⁵ TARONI/MANGIN/BÄR, FN 645, 440. TARONI/AITKEN, FN 662, 299, haben die Angabe der Belastungswahrscheinlichkeit bereits 1998 als „*arbitraire et trompeuse*“ bezeichnet.

- 509 TARONI, MANGIN und BÄR plädieren in einem 1999 in der Schweizerischen Zeitschrift für Strafrecht veröffentlichten „Brief an die Rechtsprechenden“ dafür, dass Sachverständige nicht ausschliesslich die Merkmalswahrscheinlichkeit angeben sollten. Die Merkmalswahrscheinlichkeit beantworte im Prinzip die Frage des Gerichts nicht und stifte Verwirrung; sie weisen insbesondere darauf hin, dass Ankläger und Verteidiger die Merkmalswahrscheinlichkeit für irreführende Argumente verwenden können.⁸¹⁶ TARONI et al. plädieren dafür, dass der Gutachter den Likelihood-Quotienten, die abstrakte Beweiskraft, der Spur (sei es nun eine Faser, DNA oder andere Spur) angibt.⁸¹⁷
- 510 Grundsätzlich kann ich mich diesem Vorschlag nur anschliessen; er ist definitiv eine Verbesserung gegenüber dem Status quo. Um den Zähler des Likelihood-Quotienten zu berechnen, d.h. die Wahrscheinlichkeit, dass die Spur übereinstimmt, wenn der Verdächtige der Täter ist ($P(I | \text{Schuldig})$), muss der Sachverständige die Wahrscheinlichkeit abschätzen, mit welcher diese Spur vom Spurengeber am Tatort zurückgelassen worden ist, wenn er der Täter ist. Dazu muss der Sachverständige beispielsweise wissen, ob der Verdächtige Zugang zum Tatort hatte, ohne dass ihn dies belasten würde, wie heftig die Aggression war, die zur Spur führte oder wie viel Zeit verstrich zwischen der Tat und der Sicherung der Spur.⁸¹⁸ Um den Nenner des Likelihood-Quotienten zu berechnen ($P(I | \text{Unschuldig})$), muss der Gutachter neben der Häufigkeit des DNA-Profiles in der Gruppe der möglichen Aggressoren auch die Häufigkeit des Auftretens von Spuren dieser Art am Tatort, aber ohne Bezug zur Tat, kennen.
- 511 Nimmt man den Vorschlag der Angabe des Likelihood-Quotienten ernst, so übernimmt der Sachverständige daher einen Grossteil der Beweiswürdigung. Die Frage, ob Spuren der gefundenen Art am Tatort auch ohne Bezug zur Tat vorgefunden werden können, lässt sich beispielsweise nicht beantworten, ohne dass man die Aussagen des Verdächtigen kennt. Der Vorschlag von TARONI/MANGIN/BÄR bedingt daher, dass der Gutachter das gesamte oder annähernd das gesamte Untersuchungsergebnis kennt. Meines Erachtens ist es zweifelhaft, ob es (i) effizient ist, wenn der Gutachter diesen Aufwand betreiben muss und (ii) sich mit der Aufgabenteilung zwischen Gericht und Sachverständigen vereinbaren lässt. Es besteht die Gefahr, dass damit die Beweiswürdigung delegiert wird, was aus prozessrechtlichen Gründen bedenklich ist. Genauso wenig, wie sich ein Gutachter zu Rechtsfragen äussern soll (BGE 125 II 541, 549), sollte er dem Gericht die Aufgabe der Beweiswürdigung abnehmen. Der Gutachter muss der RichterIn das Beweismittel, d.h. die Merkmalswahrscheinlichkeit, liefern. Die Würdigung dieses Beweismittels ist die ureigene Aufgabe der RichterIn.⁸¹⁹
- 512 Gibt der Sachverständige gemäss dem Vorschlag von TARONI/MANGIN/BÄR einen Likelihood-Quotienten an, so muss er auf jeden Fall begründen, wie er zu den Wahrscheinlichkeiten in Nenner und Zähler gelangt ist. Diese Begründung wird in komplexen Fällen schnell einmal Ähnlichkeiten mit der Begründung der Schuld in einem gerichtlichen

⁸¹⁶ TARONI/MANGIN/BÄR, FN 645, 440 f.

⁸¹⁷ TARONI/MANGIN/BÄR, FN 645, 440 f.

⁸¹⁸ TARONI/MANGIN/BÄR, FN 645, 441 f.

⁸¹⁹ BENDER/NACK, FN 605, Rz. 424; KAYE/KOEHLER, FN 676, 648, mit zahlreichen Hinweisen.

Urteil haben, was die Überlegung stützt, dass der Gutachter damit Aufgaben des Gerichts übernimmt.

B. Statistische Informationen sollten in natürlichen Häufigkeiten präsentiert werden

- 513 Gutachter präsentieren die Merkmalswahrscheinlichkeit oft in der Form der bedingten Wahrscheinlichkeit $P(\text{Übereinstimmung} \mid \text{Unschuldig})$; d.h. sie geben an, mit welcher Wahrscheinlichkeit eine Übereinstimmung vorliegt, wenn der Verdächtige unschuldig ist.⁸²⁰ Für das Gericht wichtiger aber ist die Wahrscheinlichkeit, dass der Angeklagte unschuldig ist, obwohl die Spur übereinstimmt, also $P(\text{Unschuldig} \mid \text{Übereinstimmung})$.
- 514 Bei der Präsentation von statistischen Daten in der Form von bedingten Wahrscheinlichkeiten besteht wie erwähnt die Gefahr, dass die beiden – semantisch ähnlichen – bedingten Wahrscheinlichkeiten verwechselt werden (*inverse fallacy*).⁸²¹ Diese Gefahr lässt sich reduzieren, wenn die statistischen Informationen nicht in der Form von Wahrscheinlichkeiten oder Prozentwerten präsentiert werden, sondern in der Form natürlicher Häufigkeiten (z.B. „3 von 20 Taxis gehören der Gesellschaft B“ statt „15 % der Taxis gehören der Gesellschaft B“).⁸²²
- 515 Die Lösung der „Alkotest-Frage“ in der Form, wie sie von GIGERENZER und Kollegen vorgeschlagen wird, zeigt, wie viel einfacher es wird, die richtigen Schlüsse aus den statistischen Angaben zu ziehen, wenn diese in der Form von natürlichen Häufigkeiten angegeben werden.⁸²³ Bei der Angabe der natürlichen Häufigkeiten lassen sich einfach Häufigkeitsbäume gemäss der Abbildung 10 zeichnen, die den Zusammenhang zwischen Anfangswahrscheinlichkeiten, falsch positiven Resultaten und Belastungswahrscheinlichkeit unmittelbar klar machen.
- 516 Von 10'000 kontrollierten Fahrern sind 100 (1 %) betrunken. Bei 95 davon ist der Alkoholtest positiv. Bei 5 %, oder 495, der 9'900 nüchternen Fahrer ist der Test aber ebenfalls positiv. Der Anteil betrunkenen Fahrer an der Gesamtzahl der positiv getesteten Fahrer ergibt sich, wenn man die Anzahl richtig positiver Testresultate durch die Anzahl richtiger positiver und falsch positiver Testresultate teilt. Entscheidend sind also letztlich nur die Zahlen in den fett umrandeten Boxen. Die Gefahr von Trugschlüssen lässt es angezeigt erscheinen, insbesondere statistische Informationen in Häufigkeiten darzustellen.

⁸²⁰ BENDER/NACK, FN 605, Rz. 421.

⁸²¹ THOMPSON/SCHUMANN, FN 621, 174.

⁸²² STEFAN KRAUS/RALPH HERTWIG, Muss DNA-Evidenz schwer verständlich sein?, Monatsschrift für Kriminologie und Strafrechtsreform 2000, 155-162, 158 f; SAMUEL LINDSEY/RALPH HERTWIG/GERD GIGERENZER, Communicating Statistical DNA Evidence, Jurimetrics 2003, 147-163, 154.

⁸²³ PETER SEDLMEIER/GERD GIGERENZER, Teaching Bayesian Reasoning in Less Than Two Hours, Journal of Experimental Psychology: General 2001, 380-400; HASTIE/DAWES, FN 399, 194. Für weitere grafischen Darstellungsweisen siehe GIGERENZER, FN 607, 108 und 113.

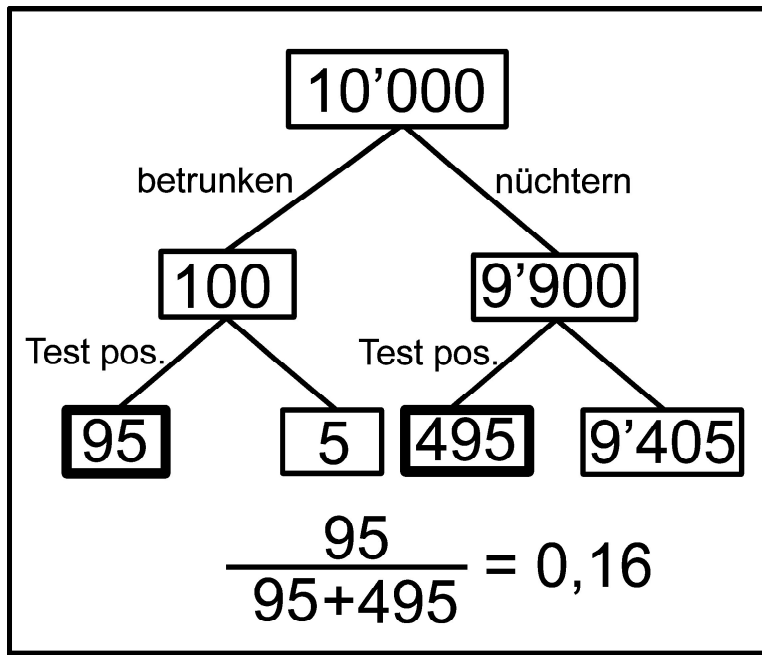


Abbildung 10: Häufigkeitsbaum des Alkotest-Problems

- 517 LINDSEY et al. haben in einem Experiment mit 127 Jurastudierenden und 27 wissenschaftlichen Angestellten der juristischen Fakultät der Freien Universität Berlin empirisch untersucht, wie sich die unterschiedliche Darstellung der statistischen Information in einem DNA-Gutachten auf die Beurteilung der Schuld des Angeklagten auswirkt.⁸²⁴ Jeder Teilnehmer der Studie erhielt detaillierte Fallbeschreibungen zweier Mordfälle, die jeweils ein DNA-Gutachten umfassten. Die Juristen hatten eine Stunde Zeit, sich in die zwei Fälle einzulesen, und sollten danach verschiedene Fragen zu den DNA-Gutachten beantworten (z.B. „Mit welcher Wahrscheinlichkeit ist der Angeklagte der Verursacher der am Tatort sichergestellten DNA-Spur?“). Auf Grundlage der verfügbaren Beweismittel (im Wesentlichen nur das DNA-Gutachten) sollten die Versuchspersonen anschliessend ihr Urteil über Schuld oder Unschuld des Angeklagten abgeben.⁸²⁵
- 518 Wurden die Informationen in Wahrscheinlichkeiten dargestellt, konnten nur ca. 8 % der Versuchspersonen die gefragten Wahrscheinlichkeiten korrekt angeben, gegenüber 40 % richtigen Antworten bei der Darstellung in natürlichen Häufigkeiten.⁸²⁶ Die wissenschaftlichen Mitarbeiter schnitten klar besser ab als die Studierenden; rund 70 % von ihnen gaben bei der Darstellung der Information in natürlichen Häufigkeiten die korrekte Antwort.⁸²⁷ Monetäre Anreize, d.h. ein Geldpreis für die richtige Antwort, hatten dabei keinen Einfluss auf die Anzahl der richtigen Antworten auf die Fragen nach der Wahrscheinlichkeit.⁸²⁸ Die natürlichen Häufigkeiten beeinflussten auch das Urteil über die Schuld: bei der Darstellung

⁸²⁴ LINDSEY/HERTWIG/GIGERENZER, FN 822.

⁸²⁵ LINDSEY/HERTWIG/GIGERENZER, FN 822, 158. Die Studie wird auch besprochen in KRAUSS/HERTWIG, FN 822, 160 f.

⁸²⁶ KRAUSS/HERTWIG, FN 822, 160.

⁸²⁷ LINDSEY/HERTWIG/GIGERENZER, FN 822, 159.

⁸²⁸ LINDSEY/HERTWIG/GIGERENZER, FN 822, 159.

in Wahrscheinlichkeiten waren mehr als die Hälfte (55 %) der Versuchspersonen von der Schuld des Angeklagten überzeugt,⁸²⁹ bei der Darstellung in natürlichen Häufigkeiten bei ansonsten gleicher Sachlage nur 33 %.⁸³⁰

- 519 Gutachter sollten ihre Angaben zur Merkmalswahrscheinlichkeit in DNA-, Faser- und ähnlichen Gutachten daher in der Form von natürlichen Häufigkeiten machen; also beispielsweise „Es kommen grundsätzlich zehn Millionen Männer als Täter in Frage. Ungefähr zehn von ihnen haben ein DNA-Profil, das mit der Tatortspur übereinstimmt. Wenn jemand dieses Profil aufweist, ist es praktisch sicher, dass die DNA-Analyse eine Übereinstimmung anzeigt. Unter den 9'999'990 Männern, deren DNA-Profil nicht mit der Tatortspur übereinstimmt, wird die DNA-Analyse in 100 Fällen ebenfalls eine Übereinstimmung feststellen“.⁸³¹ Durch diese Art der Präsentation statistischer Informationen lassen sich viele unnötige Fehler vermeiden.

VII. Beweiswürdigung und DNA-Analysen

A. Die Technik der DNA-Analyse

- 520 Die Technik der DNA⁸³²-Analyse für forensische Zwecke ist, obwohl heute Routine, sehr komplex. Da für das im Folgenden entwickelte statistische Argument kein vertieftes Verständnis des Ablaufs einer DNA-Analyse notwendig ist, wird die Technik hier nur kurz erläutert, ohne auf Details einzugehen. Vertiefte Darstellungen finden sich bei MARC HOTHER, BÄR/KRATZER und BRODERSEN/ANSLINGER/ROLF.⁸³³
- 521 Die Erbsubstanz DNA ist der chemische Stoff, der die menschliche Erbinformation enthält. Die DNA befindet sich als fadenförmiges Molekül im Kern jeder Zelle des menschlichen Körpers. Nur einige wenige Prozente des DNA-Moleküls enthalten die heute bekannten Erbfaktoren, die Gene. Die DNA verschiedener Menschen stimmt zu rund 99 % überein. Abweichungen gibt es nur an relativ wenigen, bekannten Stellen („Loci“).
- 522 Die DNA-Analyse zu Identifizierungszwecken untersucht bestimmte solcher Loci. Jeder dieser Abschnitte kann mehrere Kombinationen der Bauelemente der DANN aufweisen, so dass für einen bestimmten Abschnitt nur zehn Prozent der Menschen die gleiche Kombination aufweisen. In allen Staaten, welche die DNA-Analyse zur Identifikation einsetzen, wird von denselben untersuchten Abschnitten ausgegangen. Die Wahrscheinlichkeit, dass

⁸²⁹ Obwohl 92 % von ihnen die statistischen Informationen nicht verstanden hatten!

⁸³⁰ KRAUSS/HERTWIG, FN 822, 160.

⁸³¹ Nach LINDSEY/HERTWIG/GIGERENZER, FN 822, 163. Das Beispiel geht von einer falsch-positiv Rate von 1:100'000 aus; d.h. in einem von hunderttausend Fällen, in denen *keine* Übereinstimmung von Tatort- und Angeschuldigtenprofil vorliegt, wird *dennoch* eine Übereinstimmung festgestellt. Zur Bedeutung falsch positiver Ergebnisse gleich nachstehend.

⁸³² Desoxyribo-Nucleic-Acid, deutsch Desoxyribo-Nuklein-Säure, DNS. Die Abkürzung DNA hat sich auch im deutschsprachigen Raum durchgesetzt.

⁸³³ MARC HOTHER, Die DNA-Analyse – ihre Bedeutung für die Strafverfolgung und ihr Beweiswert im Strafverfahren, Diss. Giessen 1995; WALTER BÄR/ADELGUNDE KRATZER, Die Anwendung der Gentechnologie in der Rechtsmedizin, in: Forschung für Leben Nr. 50, Mai 1998; KILIAN BRODERSEN/KATJA ANSLINGER/BURKHARD ROLF, DNA-Analyse und Strafverfahren, München 2003.

sich zufällig zweimal das gleiche DNA-Profil vorfindet, ist umso kleiner ist, je grösser die Anzahl untersuchter Abschnitte ist, wenn man davon ausgeht, dass die Wahrscheinlichkeit der Übereinstimmung in einem bestimmten Abschnitt unabhängig von der Übereinstimmung in einem anderen Abschnitt ist (und daher die Produktregel Anwendung findet). Bei einem DNA-Profil, welches aufgrund der Analyse von 11 DNA-Abschnitten erstellt wurde, ist die Wahrscheinlichkeit kleiner als eins zu 10 Milliarden, dass ein zufällig ausgewähltes Individuum, das mit dem Probengeber der ersten Probe nicht verwandt ist, ein gleiches Profil aufweist.⁸³⁴

B. Der Beweiswert von DNA-Analysen und falsch positive Resultate

- 523 Der Beweiswert einer DNA-Analyse lässt sich mittels des Bayes-Theorems analysieren. Das Gericht oder die Strafverfolgungsbehörde möchte wissen, welche der beiden folgenden Hypothesen zutrifft:

U = die untersuchte Probe stammt vom Verdächtigen;

$\neg U$ = die untersuchte Probe stammt nicht vom Verdächtigen.

- 524 Es sei B der Bericht des Gutachters, der eine Übereinstimmung des DNA-Profiles des Verdächtigen und der Probe, die am Tatort erhoben wurde, feststellt. Immer dann, wenn eine Übereinstimmung tatsächlich vorliegt, stellt der Gutachter sie auch fest, es gilt also $P(B | U) = 1$. Der Gutachter wird aber auch dann eine Übereinstimmung feststellen, wenn die Probe nicht vom Verdächtigen stammt, aber der Verdächtige und der Spurenverursacher zufälligerweise genau den gleichen genetischen Fingerabdruck aufweisen. Diese Wahrscheinlichkeit einer zufälligen Übereinstimmung (*random match probability*, RMP) ist, wie oben gesehen, sehr gering. Es ist also $P(B | \neg U) = \text{RMP}$. Der Gutachter wird schliesslich auch dann von einer Übereinstimmung berichten, wenn er irrtümlicherweise eine Übereinstimmung feststellt, obwohl diese gar nicht vorliegt. Dieser Fehler wird „falsch positiv“ genannt, d.h. der Gutachter bejaht eine Hypothese, obwohl die Daten sie nicht belegen (*false positive probability*, FPP). Der Gutachter begeht dann keinen Fehler, wenn er eine Übereinstimmung feststellt, obwohl Verdächtiger und Spurengeber nicht identisch sind, wenn die zu vergleichenden DNA-Profile zufällig übereinstimmen. Die Wahrscheinlichkeit eines falsch positiven Befundes ist daher $\text{FPP} \cdot (1 - \text{RMP})$. Die Gesamtwahrscheinlichkeit, dass der Gutachter eine Übereinstimmung feststellt, obwohl keine Übereinstimmung vorliegt, ist folglich $P(B | \neg U) = \text{RMP} + (\text{FPP} \cdot (1 - \text{RMP}))$.
- 525 Die Wahrscheinlichkeit, dass das DNA-Profil vom Verdächtigen stammt, wenn der Gutachter von einer Übereinstimmung berichtet ist $P(U | B)$. $P(U | B)$ errechnet sich gemäss dem Bayes-Theorem in seiner extensiven Form für zwei sich gegenseitig ausschliessende Hypothesen U und $\neg U$ gemäss Gleichung 9, S. 125

⁸³⁴ Botschaft zum Bundesgesetz über die Verwendung von DNA-Profilen im Strafverfahren und zur Identifizierung von unbekannten und vermissten Personen, 37 (www.admin.ch/ch/d/ff/2001/29.pdf, besucht am 17. Januar 2005).

$$P(U | B) = \frac{P(U) \cdot P(B | U)}{P(U) \cdot P(B | U) + P(\neg U) \cdot P(B | \neg U)} \quad (21)$$

526 Ersetzt man in dieser Gleichung $P(B | \neg U)$ durch $RMP + (FPP \cdot (1 - RMP))$, so erhält man

$$P(U | B) = \frac{P(U) \cdot P(B | U)}{P(U) \cdot P(B | U) + P(\neg U) \cdot (RMP + (FPP \cdot (1 - RMP)))} \quad (22)$$

527 Als Likelihood-Quotient lässt sich die abstrakte Beweiskraft eines DNA-Gutachtens wie folgt darstellen:⁸³⁵

$$\frac{P(B | U)}{P(B | \neg U)} = \frac{1}{RMP + (FPP \cdot (1 - RMP))} \quad (23)$$

528 In dieser Form der Darstellung sieht man besonders klar, dass der Likelihood-Quotient bei sehr kleiner RMP fast ausschliesslich von der FPP bestimmt wird; es gilt also bei sehr kleiner RMP approximativ: abstrakte Beweiskraft $\approx 1/FPP$.

529 Die Diskussion der Zuverlässigkeit von DNA-Fingerabdrücken als Beweismittel hat sich bisher hauptsächlich darum gedreht, wie hoch die Wahrscheinlichkeit einer zufälligen Übereinstimmung vom DNA-Profil des Verdächtigen und dem Profil der Tatortspur ist. Es wurde befürchtet, dass diese sehr viel grösser ist, wenn Verdächtiger und Spurengerber der gleichen Ethnie angehören, weil die Verteilung der DNA-Polymorphismen möglicherweise nicht homogen ist.⁸³⁶ Auch wurde anfänglich bezweifelt, dass die Übereinstimmungen in verschiedenen Loci unabhängig sind und man die Produktregel anwenden darf, um die Wahrscheinlichkeit der gleichzeitigen Übereinstimmung mehrerer Abschnitte zu berechnen.⁸³⁷ Diese Befürchtungen haben sich jedoch nicht bestätigt. Die Frage, welche Rolle die Wahrscheinlichkeit eines falsch positiven Befundes für den Beweiswert eines DNA-Gutachtens spielt, wurde bisher kaum diskutiert.⁸³⁸ Dabei kann deren Einfluss sehr gross sein, wie nachfolgendes Beispiel zeigt.

530 Angenommen, die Strafverfolgungsbehörde geht von einer Wahrscheinlichkeit von 10 % aus, dass die untersuchte Probe vom Verdächtigen stammt, die Wahrscheinlichkeit einer zufälligen Übereinstimmung (RMP) sei 1:1 Milliarde (die Wahrscheinlichkeit entsprechend 10^{-9}) und die Behörde hält es für ausgeschlossen, dass der Gutachter einen Fehler

⁸³⁵ Formaler Beweis bei THOMPSON/TARONI/AITKEN, FN 670, 53 f.

⁸³⁶ DAVID H. KAYE, DNA Evidence: Probability, Population Genetics, and the Courts, Harvard Journal of Law & Technology 1993, 101-172, 128 ff.

⁸³⁷ BGHSt 38, 320.

⁸³⁸ Siehe jedoch JONATHAN J. KOEHLER/AUDREY CHIA/J. SAM LINDSEY, The Random Match Probability (RMP) in DNA Evidence: Irrelevant and Prejudicial?, Jurimetrics 1995, 201-219; JONATHAN J. KOEHLER, Why DNA Likelihood-Ratios Should Account for Error (Even When A National Research Council Report Says They Should Not), Jurimetrics 1997, 425-437.

Besonderer Teil

macht und eine Übereinstimmung meldet, obwohl keine vorliegt (FPP = 0). Dann gilt numerisch

$$P(U | B) = \frac{1 \cdot 0,1}{1 \cdot 0,1 + 0,9 \cdot 10^{-9} + (0 \cdot (1-10^{-9}))} = 0.999999991$$

531 Die Chance, dass die untersuchte Probe vom Verdächtigen stammt, beträgt demnach 100'000'000:1, es ist mit anderen Worten praktisch sicher, dass der Verdächtige der Spurengerber ist (ob er der Täter ist, ist wieder eine andere Frage – möglicherweise ist seine DNA an den Tatort gelangt, ohne dass er der Täter ist).

532 Was aber, wenn eine geringe Wahrscheinlichkeit von einem Tausendstel besteht, dass der Gutachter ein falsch positives Resultat meldet?

533 Es gilt dann numerisch

$$P(U | B) = \frac{1 \cdot 0,1}{1 \cdot 0,1 + 0,9 \cdot 10^{-9} + (0,001 \cdot (1-10^{-9}))} = 0.99$$

534 Die Chance, dass der Verdächtige der Spurengerber ist, beträgt demnach nur noch 100 zu 1 und ist um Größenordnungen kleiner als bei einer FPP von 0. Man darf daher die Gefahr, dass ein falsch positives Resultat gemeldet wird, bei der Beurteilung des Beweiswerts von DNA-Gutachten auf keinen Fall vernachlässigen. Die nachstehend abgebildete Tabelle 13 aus THOMPSON/TARONI/AITKEN zeigt, wie sich die nachträgliche Chance (*posterior odds*) im Verhältnis zur anfänglichen Chance, Wahrscheinlichkeit zufälliger Übereinstimmung und nachträglicher Wahrscheinlichkeit ändert. Wie leicht ersichtlich, ist die hohe Ausschlussleistung der DNA-Analyse („Chance einer zufälligen Übereinstimmung von eins zu zehn Milliarden“) irreführend, denn die Belastungswahrscheinlichkeit hängt weitgehend von der Wahrscheinlichkeit eines falsch positiven Fehlers ab.

Anfängliche Chance (Prior Odds)	Wahrscheinlichkeit zufälliger Übereinstimmung (RMP)	Wahrscheinlichkeit eines falsch positiven Resultates (FPP)	Nachträgliche Wahrscheinlichkeit (Posterior Odds)
2:1	10 ⁻⁹	0	2'000'000'000
2:1	10 ⁻⁹	0.0001	20'000
2:1	10 ⁻⁹	0.001	2'000
2:1	10 ⁻⁹	0.01	200
2:1	10 ⁻⁶	0	2'000'000
2:1	10 ⁻⁶	0.0001	19'802
2:1	10 ⁻⁶	0.001	1'998
2:1	10 ⁻⁶	0.01	200
2:1	10 ⁻³	0	2'000
2:1	10 ⁻³	0.0001	1'818
2:1	10 ⁻³	0.001	1'001
2:1	10 ⁻³	0.01	182
1:10	10 ⁻⁹	0	100'000'000
1:10	10 ⁻⁹	0.0001	1'000

1:10	10^{-9}	0.001	100
1:10	10^{-9}	0.01	10
1:10	10^{-6}	0	100'000
1:10	10^{-6}	0.0001	990
1:10	10^{-6}	0.001	100
1:10	10^{-6}	0.01	10
1:10	10^{-3}	0	100
1:10	10^{-3}	0.0001	91
1:10	10^{-3}	0.001	50
1:10	10^{-3}	0.01	9
1:100	10^{-9}	0	10'000'000
1:100	10^{-9}	0.0001	100
1:100	10^{-9}	0.001	10
1:100	10^{-9}	0.01	1
1:100	10^{-6}	0	10'000
1:100	10^{-6}	0.0001	99
1:100	10^{-6}	0.001	10
1:100	10^{-6}	0.01	1
1:100	10^{-3}	0	10
1:100	10^{-3}	0.0001	9
1:100	10^{-3}	0.001	5
1:100	10^{-3}	0.01	1
1:1'000	10^{-9}	0	1'000'000
1:1'000	10^{-9}	0.0001	10
1:1'000	10^{-9}	0.001	1
1:1'000	10^{-9}	0.01	0,1
1:1'000	10^{-6}	0	1'000
1:1'000	10^{-6}	0.0001	9,9
1:1'000	10^{-6}	0.001	1
1:1'000	10^{-6}	0.01	0,1
1:1'000	10^{-3}	0	1
1:1'000	10^{-3}	0.0001	0,91
1:1'000	10^{-3}	0.001	0,50
1:1'000	10^{-3}	0.01	0,09

Tabelle 13: Beweiswert eines DNA-Gutachtens in Abhängigkeit von Fehlerwahrscheinlichkeit⁸³⁹

- 535 Die geringe Anfangschance von 1:1'000 kann sich ergeben, wenn es keine belastenden Indizien ausser der Übereinstimmung der DNA-Proben gibt, wie dies beispielsweise der Fall sein kann, wenn der Treffer durch die Suche in einer Datenbank mit den DNA-Profilen von tausenden von Personen gefunden wurde.⁸⁴⁰
- 536 THOMPSON/TARONI/AITKEN weisen darauf hin, dass die Vorstellung, die falsch positive Rate sei unerheblich, auf einem ähnlichen Denkfehler wie der Trugschluss des Anklägers beruht; sie sprechen von *false positive fallacy*.⁸⁴¹ Man nimmt irrtümlicherweise an, dass, wenn die falsch positive Rate gering ist, die Wahrscheinlichkeit einer zu Unrecht gefunde-

⁸³⁹ Aus THOMPSON/TARONI/AITKEN, FN 670, 50.

⁸⁴⁰ THOMPSON/TARONI/AITKEN, FN 670, 50.

⁸⁴¹ THOMPSON/TARONI/AITKEN, FN 670, 51.

nen Übereinstimmung ebenfalls gering ist. Mit anderen Worten, wenn die Wahrscheinlichkeit, dass die Übereinstimmung zu Unrecht gefunden wurde, 1 % beträgt, dann „muss“ folglich die Wahrscheinlichkeit, dass die Übereinstimmung zu Recht gefunden wurde, 99 % betragen. Diese Überlegung ist, wie vorne bereits diskutiert, falsch, denn die Wahrscheinlichkeit eines falsch positiven Resultats in einem Einzelfall kann nicht ohne Berücksichtigung der Anfangswahrscheinlichkeit, dass die Proben übereinstimmen, berechnet werden.

537 Auch hier mag ein Häufigkeitsbaum die Analyse veranschaulichen.

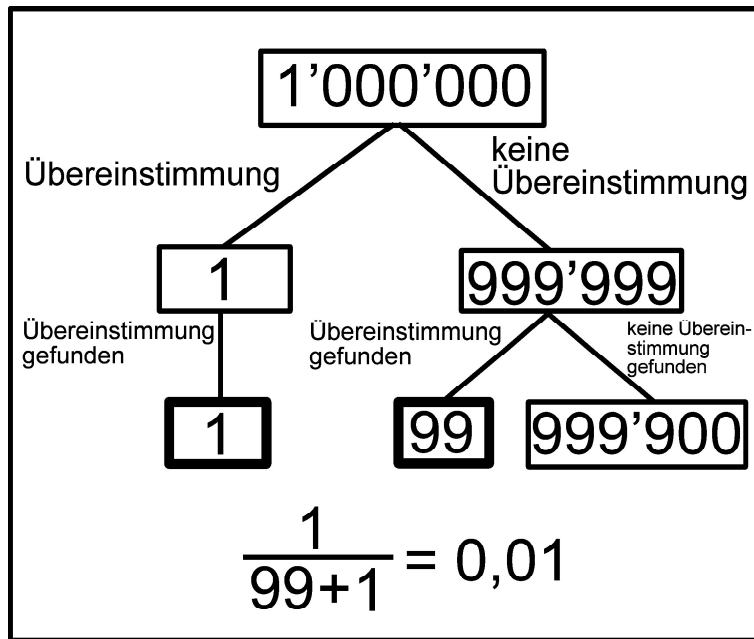


Abbildung 11: Häufigkeitsbaum bei DNA-Analyse

538 Das Abbildung 11 dargestellte Beispiel geht davon aus, dass die Anfangswahrscheinlichkeit bei 1:1 Million liegt und die RMP ebenfalls 1:1 Million beträgt. Beim Spurengeber – es kann nur einen geben – wird das Gutachten eine Übereinstimmung finden. Bei den 999'999 Verdächtigen, die nicht Spurengeber sind, wird das Gutachten aber bei einer falsch positiven Rate von 1:10'000, wie im Beispiel angenommen, ebenfalls bei 99 Personen fälschlicherweise eine Übereinstimmung feststellen, obwohl keine vorliegt. Die Wahrscheinlichkeit, dass eine Person, deren DNA-Profil mit der Tatortspur übereinstimmt, tatsächlich der Täter ist, ist daher $1/(99 + 1) = 0,01$.

539 Die vorgehenden Überlegungen werfen natürlich die Frage auf, wie hoch denn die falsch positiv Rate in Wirklichkeit ist. Dies ist eine empirisch nur unzulänglich erforschte Frage. Als forensische DNA-Analysen erstmals vor Gericht zugelassen wurden, gab es eine Reihe von Experten, die behaupteten, Fehler kämen bei der DNA-Analyse gar nie vor.⁸⁴² Diese Behauptung wurde inzwischen in Theorie und Praxis widerlegt. Fehler können bei der Entnahme und der Handhabung von Proben vorkommen oder auf eine Fehlinterpretation oder Fehlübermittlung der Testresultate zurückzuführen sein. Besonders bei Proben, die

⁸⁴² Nachweise bei THOMPSON/TARONI/AITKEN, FN 670, 47.

am Tatort erhoben werden und Spuren verschiedener DNA enthalten, ist die Separation der einzelnen DNA-Stränge mitunter schwierig. Das Abschneiden eines Labors bei der Analyse von reinen Proben, wie sie in Probeläufen und bei Abstammungsgutachten verwendet werden, sagt daher wenig darüber aus, wie hoch die Wahrscheinlichkeit eines falsch positiven Befundes bei einer verunreinigten Spur ist. Dennoch wurden in *proficiency tests*, wie sie bei amerikanischen Labors regelmässig vorgenommen werden, wiederholt falsch positive Resultate gemeldet.⁸⁴³ In mindestens einem Fall hat ein Laborfehler zur Verurteilung eines Unschuldigen geführt.⁸⁴⁴ Für die frühen 90-er Jahre des letzten Jahrhunderts wurden Fehlerraten von 0,0008 bis 0,04 geschätzt.⁸⁴⁵ Cellmark Ltd, der führende Anbieter von forensischen DNA-Tests in den USA, musste auf Druck der Verteidigung im O.J. Simpson Prozess zugeben, dass seine eigene Fehlerrate bei rund 1:200 lag.⁸⁴⁶

- 540 Die beste Methode, um die falsch positiv Rate empirisch festzustellen, sind Blindtests durch unabhängige Stellen.⁸⁴⁷ In der Schweiz sieht die DNA-Profil-Verordnung (SR 363.1) vor, dass sich Labors, die forensische DNA-Gutachten erstellen, akkreditieren müssen (Art. 2 Abs. 1 DNA-Profil-Verordnung). Um akkreditiert zu werden, muss ein Labor Qualitätsanforderungen erfüllen, die vom Bund vorgegeben und in einer Vereinbarung festgehalten werden, sowie unter der Leitung einer qualifizierten Fachperson stehen (Art. 2 Abs. 2 lit. b und c DNA-Profil-Verordnung). Akkreditiert werden forensische DNA-Labors durch die Schweizerische Akkreditierungsstelle (SAS).⁸⁴⁸ Die Labors nehmen an Eignungsprüfungen teil, bei denen eine unabhängige Stelle, z.B. die Schweizerische Gesellschaft für Rechtsmedizin, den Labors biologisches Material zur Analyse zustellt und das Resultat überprüft.
- 541 Bevor ein DNA-Profil in die DNA-Datenbank des Bundes aufgenommen wird, muss es in zwei unabhängigen durchgeführten Analyseprozessen dasselbe Resultat zeitigen (Doppelanalyse); diese Anforderung ergibt sich aus den durch das EJPD im Rahmen von Art. 2 Abs. 2 lit. b DNA-Profil-Verordnungen festgelegten Leistungs- und Qualitätsanforderungen.⁸⁴⁹
- 542 Blindtests sind in der Schweiz jedoch im Rahmen des Akkreditierungs- und Qualitätssicherungsverfahrens nicht vorgesehen; d.h. die akkreditierten Labors wissen immer, wann sie geprüft werden.⁸⁵⁰

⁸⁴³ Nachweise bei THOMPSON/TARONI/AITKEN, FN 670, 47.

⁸⁴⁴ Timothy Durham, Nachweise bei THOMPSON/TARONI/AITKEN, FN 670, 48.

⁸⁴⁵ KOEHLER/CHIA/LINDSEY, FN 838, 206 ff.

⁸⁴⁶ Genau genommen bei $2:365 = 0,0055$; JONATHAN J. KOEHLER, One in Millions, Billions and Trillions: Lessons from People v. Collins (1968) for People v. Simpson (1995), Journal of Legal Education 1997, 214-223, 221.

⁸⁴⁷ THOMPSON/TARONI/AITKEN, FN 670, 53.

⁸⁴⁸ Informationen zur SAS unter www.sas.ch/de/portraet/index.html (besucht am 4. Februar 2005). Die Tätigkeit der SAS wird durch die Verordnung über das schweizerische Akkreditierungssystem und die Bezeichnung von Prüf-, Konformitätsbewertungs-, Anmelde- und Zulassungsstellen (Akkreditierungs- und Bezeichnungsverordnung, AkkBV, SR 946.512) geregelt.

⁸⁴⁹ Schriftliche Auskunft von Christian Linsi, stv. Chef Rechtsdienst des Bundesamtes für Polizei, vom 28. Januar 2005.

⁸⁵⁰ Auskunft gemäss FN 849.

- 543 Die Proben, die analysiert werden, um das in die DNA-Profil-Datenbank aufzunehmende DNA-Profil zu bestimmen, sind von guter Qualität. Sie werden unter optimalen Bedingungen mit Wangenschleimhautabstrich vom Probengeber entnommen. Die Gefahr einer Fehlanalyse ist daher geringer als bei Tatortspuren, die oft mit fremd-DNA verunreinigt sind. Ein falsch positives Resultat kann nicht nur dadurch entstehen, dass das in der DNA-Profil-Datenbank aufgenommene Profil falsch bestimmt wurde, sondern auch dadurch dass das Profil der Tatortspur falsch bestimmt wird. Die Doppelanalyse aller DNA-Proben, deren Profile in die DNA-Profil-Datenbanken aufgenommen werden, löst daher nur das geringere Problem.
- 544 Mit einer gewissen Regelmässigkeit durchgeführte Blindtests aller DNA-Labors sind zugegebenermassen aufwendig und teuer.⁸⁵¹ Wiederholt wurde deshalb vorgeschlagen, dass die Gefahr von falsch positiven Treffern mit weniger Aufwand durch Doppelanalysen vermieden werden könne.⁸⁵² Dies scheint auch die Auffassung des Bundesamtes für Polizei zu sein.
- 545 Doppelanalysen sind aber nicht immer durchführbar, weil die am Tatort erhobenen Proben unter Umständen beim ersten Test verbraucht wurden. Eine Doppelanalyse kann auch nicht jeden Fehler korrigieren – wenn eine Kontaminierung der Probe erfolgte, ehe sie geteilt wurde, nützt auch eine zweite Analyse derselben Probe nichts.⁸⁵³ Die Fehlerreduktion durch Doppelanalyse ist schliesslich nur dann optimal, wenn die Ergebnisse der beiden Analysen untereinander unabhängig sind. Dies muss aber durchaus nicht der Fall sein – wenn der Fehler beim ersten Mal darauf zurückzuführen ist, dass die Probe besonders schwierig zu analysieren oder das Profil besonders schwierig zu interpretieren war, dann wird der gleiche Fehler mit erhöhter Wahrscheinlichkeit auch bei der zweiten Analyse gemacht werden. Doppelanalysen können die Wahrscheinlichkeit eines falsch positiven Ergebnisses wohl reduzieren, aber niemand behauptet, dass sie es ausschliessen können.⁸⁵⁴
- 546 Gerichte sollten sich bewusst sein, dass DNA-Proben zweifellos sehr beweiskräftige Indizien sind. Sie sollten aber ob der spektakulären Erfolge der forensischen DNA-Analyse

⁸⁵¹ JOSEPH L. PETERSON/R. E. GAENSSLEN, Developing Criteria for Model External DNA Proficiency Testing: Final Report National Institute of Justice, 2001, xi (erhältlich unter www.uic.edu/pharmacy/depts/forensicsci/NIJ%20Blind%20DNA%20PT%20Final%20Report.pdf, besucht am 5. Februar 2005).

⁸⁵² National Research Council, The Evaluation of Forensic DNA Evidence, Washington DC 1996, 87: "A wrongly accused person's best insurance against the possibility of being falsely incriminated is the opportunity to have the testing repeated".

⁸⁵³ THOMPSON/TARONI/AITKEN, FN 670, 48.

⁸⁵⁴ THOMPSON/TARONI/AITKEN, FN 670, 48.

– z.B. im Falle des Mordes des deutschen Modeschöpfers Rudolf Mooshammer im Januar 2005 – nicht vergessen, dass die DNA-Analyse nicht unfehlbar ist. Das Verständnis des Bayes-Theorems hilft aufzeigen, welche Faktoren den abstrakten Beweiswert eines DNA-Gutachtens bestimmen. Kein Richter sollte sich aus falscher Rücksicht auf die Unfehlbarkeit der DNA-Analyse dazu gedrängt fühlen, einen Angeklagten, gegen den als einziges Beweismittel ein DNA-Gutachten vorliegt, zu verurteilen. Vernünftige Zweifel an der Schuld des Angeklagten sind in solchen Fällen durchaus rational zu begründen.

§ 12 Bestätigungsfehler (*confirmation bias*)

- 547 Unter dem Oberbegriff Bestätigungsfehler (*confirmation bias*) wird eine Vielzahl verschiedener Phänomene zusammengefasst. Dementsprechend unterschiedlich sind die Definitionen des Begriffs in der Literatur.⁸⁵⁵ Hier soll unter Bestätigungstendenz das Phänomen verstanden werden, dass Informationen, die eine von der Versuchsperson zumindest in Betracht gezogene Hypothese bestätigen, von ihr eher gesucht, wahrgenommen, stärker gewichtet oder besser in Erinnerung behalten werden als Informationen, die gegen die Hypothese sprechen.⁸⁵⁶ Noch allgemeiner formuliert liegt eine Bestätigungstendenz vor, wenn die Neigung besteht, eine zumindest in Betracht gezogene Hypothese beizubehalten und eine korrespondierende Abneigung, sie zugunsten einer neuen Hypothese aufzugeben.⁸⁵⁷
- 548 Während man sich weitgehend einig ist, dass Bestätigungstendenzen real sind,⁸⁵⁸ besteht über die Ursachen der beobachteten Phänomene bei weitem keine Einigkeit. Können Menschen Hypothesen nicht neutral und unvoreingenommen testen oder wollen sie nicht? Wenn kognitive Faktoren zumindest mitverantwortlich sind für die beobachteten Tendenzen: Sind Bestätigungstendenzen darauf zurückzuführen, dass Menschen nicht in der Lage sind, mehr als eine Alternative gleichzeitig zu beachten?⁸⁵⁹ Oder darauf, dass Menschen das Vorhandensein einer Eigenschaft leichter erkennen und erinnern können als ihr Fehlen?⁸⁶⁰ Oder kann erst eine Interaktion verschiedener Phänomene, wie einer positiven Teststrategie und der leichteren Verarbeitung positiver Informationen, Bestätigungsfehler erklären?⁸⁶¹
- 549 Nicht nur keine Einigkeit, sondern Streit herrscht bezüglich der Frage, ob Bestätigungstendenzen als „Fehler“ bezeichnet werden dürfen; d.h. ob sie „irrational“ sind und zu Resultaten führen, die suboptimal sind. Während bis Mitte der 80-er Jahre des letzten Jahrhunderts die Ansicht vorherrschte, dass Bestätigungstendenzen ein schweres Defizit menschlichen Denkens mit negativen Konsequenzen sind, hat sich seither das Bild gewandelt. Zahlreiche Autoren weisen darauf hin, dass das, was vorschnell als Bestätigungsfehler bezeichnet wurde, ausserhalb des Labors oft eine der Umwelt angepasste Strategie ist, die

⁸⁵⁵ JOSHUA KLAYMAN, Varieties of Confirmation Bias, in: JEROME BUSEMEYER/REID HASTIE/DOUGLAS L. MEDIN (Hrsg.), Decision Making From a Cognitive Perspective, New York 1995, 385-418, 385: “There are almost as many operational definitions of *confirmation bias* as there are studies”; RAYMOND S. NICKERSON, Confirmation Bias: A Ubiquitous Phenomenon in Many Guises, Review of General Psychology 1998, 175-220, 220.

⁸⁵⁶ Nach KLAUS FIEDLER, Beruhen Bestätigungsfehler nur auf einem Bestätigungsfehler?, Psychologische Beiträge 1983, 280-286, 282.

⁸⁵⁷ KLAYMAN, FN 855, 386.

⁸⁵⁸ KLAYMAN, FN 855, 385; NICKERSON, FN 855, 177; FIEDLER, FN 856, 282.

⁸⁵⁹ CLIFFORD R. MYNATT/MICHAEL E. DOHERTY/WILLIAM DRAGAN, Information Relevance, Working Memory, and the Consideration of Alternatives, The Quarterly Journal of Experimental Psychology 1993, 759-788.

⁸⁶⁰ JONATHAN ST. B. T. EVANS, Bias in Human Reasoning: Causes and Consequences, Hove 1989.

⁸⁶¹ KLAYMAN, FN 855, 399.

mit minimalem Aufwand wenn nicht zu optimalen, so doch zu vernünftigen Resultaten gelangt.⁸⁶²

- 550 Eine Bestätigungstendenz kann erstens daraus resultieren, dass man von vorneherein nur solche Informationsquellen auswählt, von denen bestätigende Informationen zu erwarten sind. Zweitens kann eine Bestätigung zu Unrecht erfolgen, wenn ambivalente Informationen, die sowohl für wie gegen die Hypothese sprechen können, konsequent als bestätigend codiert werden. Drittens kann eine unangebrachte Bestätigung erfolgen, wenn man zwar sowohl bestätigende wie nichtbestätigende Informationen registriert, letztere aber nicht (oder nur ungenügend) zur Beurteilung der Hypothese heranzieht.⁸⁶³ Nach dieser Systematik lassen sich Bestätigungstendenzen bei der Informationssuche und bei der Informationsinterpretation unterscheiden. Sie dient als Gerüst der folgenden Übersicht, die sich massgeblich auf die Übersichtsartikel von NICKERSON und KLAYMAN stützt.⁸⁶⁴ In Anbetracht der beinahe unüberschaubaren Literatur zum Hypothesentesten muss diese Übersicht notwendigerweise verkürzt und unvollständig ausfallen. Nach der Darstellung der empirischen Belege für Bestätigungstendenzen werden ihre Folgen beleuchtet, ehe die Resultate der eigenen Studie rapportiert und diskutiert werden; schliesslich wird dargestellt, was gegen den Einfluss des *confirmation bias* unternommen werden kann.

I. Informationssuche

A. Pseudodiagnostizität

- 551 Der Begriff „Pseudodiagnostizität“ wurde von DOHERTY et al. geprägt.⁸⁶⁵ Wie vorne, S. 133 ff., dargelegt wurde, beurteilt sich die abstrakte Beweiskraft eines Indizes I danach, wie viel häufiger (seltener) es bei Zutreffen der Hypothese H vorliegt als bei Nichtzutreffen von H (abstrakte Beweiskraft = $P(I | H)/P(I | \neg H)$ = Likelihood-Quotient). Die meisten Menschen scheinen dieses Konzept wie bereits erwähnt nicht intuitiv zu verstehen. In einer typischen Studie zur Pseudodiagnostizität werden die Versuchspersonen gebeten, anzugeben, welcher Kategorie eine Instanz wahrscheinlich angehört; beispielsweise von welchem Fundort ein Gefäss stammt.⁸⁶⁶ Ein Gefäss aus glattem Ton mit gebogenen Henkeln kann beispielsweise entweder von der Muschel-Insel oder von der Korallen-Insel kommen. Welche Informationen werden benötigt, um zu entscheiden, ob das Gefäss eher von der Muschel- oder von der Korallen-Insel stammt? Nachdem die Versuchspersonen sich aufgrund der bekannten Merkmale eine erste Meinung zur Herkunft des Gefässes gebildet hatten, konnten sie weitere Informationen über die Verteilung der Eigenschaften

⁸⁶² VOLKER GADENNE, Der Bestätigungsfehler und die Rationalität kognitiver Prozesse, Psychologische Beiträge 1982, 11-25; JOSHUA KLAYMAN/YOUNG-WON HA, Confirmation, Disconfirmation, and Information in Hypothesis Testing, Psychological Review 1987, 211-228; JAMES FRIEDRICH, Primary Error Detection and Minimization (PEDMIN) Strategies in Social Cognition: A Reinterpretation of Confirmation Bias Phenomena, Psychological Review 1993, 298-319.

⁸⁶³ GADENNE, FN 862, 13.

⁸⁶⁴ NICKERSON, FN 855; KLAYMAN, FN 855.

⁸⁶⁵ MICHAEL E. DOHERTY/CLIFFORD R. MYNATT/Ryan D. TWENEY/MICHAEL D. SCHIAVO, Pseudodiagnosticity, Acta Psychologica 1979, 111-121.

⁸⁶⁶ DOHERTY/MYNATT/TWENEY/SCHIAVO, FN 866, 114 f.

der Gefässe beider Inseln erwerben. Die meisten Versuchspersonen, die sich aufgrund des ersten Eindrucks für die Korallen-Insel entschieden hatten, fragten nur oder vorwiegend nach Informationen zu den Gefässen von der Korallen-Insel.⁸⁶⁷ Entscheidend ist aber immer das Verhältnis der Merkmale; d.h. die Information, dass 80 % der Gefässe von der Korallen-Insel gebogene Henkel haben, ist nur diagnostisch, wenn ich weiss, dass nicht auch 80 % der Gefässe von der Muschel-Insel gebogene Henkel haben. Wenn ich nicht danach frage, werde ich es nie erfahren. Wie vorne erwähnt, scheinen Menschen zu meinen, die Kenntnis der Typizität oder Repräsentativität einer Eigenschaft für die Kategorie genüge, um auf die Zugehörigkeit der Instanz zur Kategorie zu schliessen.⁸⁶⁸ DOHERTY et al. haben die Tendenz ihrer Versuchspersonen, sich mehr Informationen zu der Insel zu beschaffen, die sie ursprünglich favorisierten, *confirmation bias* genannt.⁸⁶⁹ Diese Bezeichnung ist nicht unumstritten.⁸⁷⁰ Richtigerweise spricht man wohl besser einfach von einer ineffizienten Suchstrategie, die weder bestätigende noch widerlegende Informationen liefern kann.

- 552 Der Anteil der nachgefragten pseudodiagnostischen Informationen sinkt, je mehr die Versuchspersonen Zweifel an der ursprünglichen Hypothese haben.⁸⁷¹ MYNATT et al. erklären dies damit, dass die Versuchspersonen eher bereit sind, eine alternative Hypothese in Betracht zu ziehen, wenn die ursprünglich Hypothese von den vorhandenen Daten nur schwach gestützt wird. Pseudodiagnostizität resultiert ihrer Meinung nach daraus, dass Menschen nicht in der Lage sind, mehr als eine Hypothese gleichzeitig zu verfolgen. Verschiedene Hypothesen werden seriell, nacheinander geprüft. Erst der Wechsel von der einen Hypothese zur anderen bewirkt, dass Informationen zu den Eigenschaften der Instanzen der anderen Kategorie abgefragt werden.⁸⁷²
- 553 FEENEY und EVANS weisen darauf hin, dass „pseudodiagnostische“ Information wertvoll sein kann, wenn das Indiz, das die ursprüngliche Hypothese stützt, sehr häufig ist.⁸⁷³ Angenommen, ich weiss nicht mehr, ob meine Schwester ein Auto der Marke X oder Y hat. Ich weiss aber, dass das Auto ein Radio hat und eine Höchstgeschwindigkeit von 280 km/h erreicht. 80 % der Autos der Marke X haben ein Radio. Streng logisch müsste ich mich danach erkundigen, wie viele Autos der Marke Y ebenfalls ein Radio haben. Da aber die meisten Autos heute ein Radio haben, wird die Antwort auf diese Frage nicht viel weiterhelfen. Die Frage danach, wie viele der Autos der Marke X eine Höchstgeschwindigkeit von mehr 280 km/h erreichen, ist zwar streng genommen pseudodiagnostisch; wenn ich aber erfahre, dass sehr viele Autos der Marke X so schnell sind, potentiell wertvoller als

⁸⁶⁷ DOHERTY/MYNATT/TWENEY/SCHIAVO, FN 866, 117.

⁸⁶⁸ KLAYMAN, FN 855, 397.

⁸⁶⁹ DOHERTY/MYNATT/TWENEY/SCHIAVO, FN 866, 116.

⁸⁷⁰ DAWES, FN 652, 12088.

⁸⁷¹ MYNATT/DOHERTY/DRAGAN, FN 859, 774; JONATHAN ST.B.T. EVANS/SIMON VENN/AIDAN FEENEY, *Implicit and Explicit Processes in a Hypothesis Testing Task*, *British Journal of Psychology* 2002, 31-46.

⁸⁷² MYNATT/DOHERTY/DRAGAN, FN 859, 774 f.

⁸⁷³ AIDAN FEENEY/JONATHAN ST.B.T. EVANS/SIMON VENN, *A Rarity Heuristic for Hypothesis Testing*, in: LILA R. GLEITMAN/ARAVID K. JOSHI (Hrsg.), *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*, Mahwah 2000, 119-124; ob eine Eigenschaft selten ist, können die Versuchspersonen auch aufgrund ihres Hintergrundwissens ahnen, AIDAN FEENEY/JONATHAN ST.B.T. EVANS/JOHN CLIBBENS, *Background Beliefs and Evidence Interpretation*, *Thinking and Reasoning* 2000, 97-124.

die Häufigkeit von Autoradios bei Y.⁸⁷⁴ FEENEY et al. können nachweisen, dass Menschen bei der Suche nach Informationen berücksichtigen, wie wertvoll die gesuchte Information potenziell sein kann. Ist das bekannte Indiz selten, werden weniger häufig pseudodiagnostische Informationen gesucht als wenn das bekannte Indiz häufig ist; wie erwähnt ist die Suche nach pseudodiagnostischer Information im zweiten Fall nicht von vorneherein irrational.

B. Bevorzugte Suche nach bestätigenden Informationen

- 554 Eines der klassischen Experimente, das zur Begründung der These, dass Menschen nicht nach falsifizierenden Informationen suchen, herangezogen wird, stammt von PETER WASON.⁸⁷⁵ Den Versuchspersonen wird ein Zahlentripel 2, 4, 6 vorgelegt. Sie erhalten die Aufgabe, die Regel herauszufinden, nach der die Folge gebildet wurde. Hierzu sollen sie weitere Zahlentripel bilden, die ihrer Meinung nach der Regel entsprechen. Nach jedem Versuch erhalten sie vom Versuchsleiter eine Rückmeldung darüber, ob ihre produzierte Folge zu der wirklichen Regel passt oder nicht. Die Versuchspersonen dürfen ihre produzierten Folgen, ihre Hypothesen und die Antworten des Versuchsleiters schriftlich festhalten. Den Versuchspersonen wird weiter erklärt, dass sie die gesuchte Regel erst dann als Ergebnis bekannt geben sollen, wenn sie aufgrund ihrer Versuche ziemlich sicher seien, sie gefunden zu haben. Die zu findende Regel lautet, dass jede folgende Zahl grösser als die vorangehende Zahl ist.
- 555 72 % der Versuchspersonen fanden die Lösung, aber nur 21 %, ohne mindestens eine falsche Hypothese zu nennen. Die meisten Versuchspersonen bilden eine bestimmte Hypothese (z.B. „Folgende Zahl = Vorhergehende Zahl + 2“) und produzieren dann ausschliesslich positive Instanzen für diese Hypothese (d.h. Zahlentripel, die der Hypothese entsprechen, wie 10, 12, 14).⁸⁷⁶ Kaum eine Versuchsperson formuliert negative Instanzen für ihre Hypothese. WASON und JOHNSON-LAIRD ziehen daraus den Schluss, dass sich die Versuchspersonen nicht gemäss der Methode des kritischen Rationalismus (POPPER) verhalten, weil sie nur versuchen, ihre Hypothese zu verifizieren.⁸⁷⁷
- 556 Das besprochene Experiment ist aber nicht geeignet, dies zu beweisen. Wie GADENNE zu Recht anführt, ist die Strategie, positive Instanzen einer Theorie zu testen, nicht notwendigerweise eine Bestätigungstendenz. Die Versuchspersonen wissen nicht, ob sie eine positive oder negative Rückmeldung durch den Versuchsleiter erhalten werden. Jede positive Instanz einer Hypothese ist geeignet, die Hypothese zu falsifizieren.⁸⁷⁸ Dass die Versuchspersonen vorwiegend bestätigendes Feedback erhalten haben, liegt daran, dass die von den meisten formulierte Theorie „Folgende Zahl = Vorhergehende Zahl + 2“ enger ist als die gesuchte Theorie „ansteigende Zahlen“. Die Vorgabe des Zahlentripels 2, 4, 6 hat

⁸⁷⁴ FEENEY/EVANS/VENN, FN 873, 121.

⁸⁷⁵ PETER C. WASON, On the Failure to Eliminate Hypothesis in a Conceptual Task, *Quarterly Journal of Experimental Psychology* 1960, 129-140.

⁸⁷⁶ WASON, FN 875, 135.

⁸⁷⁷ PETER C. WASON/PHILIP N. JOHNSON-LAIRD, *Psychology of Reasoning*, London 1972, 210.

⁸⁷⁸ GADENNE, FN 862, 18 f.

sicher dazu beigetragen, dass die Versuchspersonen eine enge Hypothese entwickelten. Die Wahrscheinlichkeit, dass ein nach der Regel „Folgende Zahl = Vorhergehende Zahl + 2“ gebildetes Zahlentripel aus der Menge der Zahlen zwischen 0 und 1'000 gezogen wird, wenn die Regel der Ziehung „ansteigende Zahlen“ lautet, ist sehr gering im Vergleich zur Wahrscheinlichkeit, dass ein beliebiges ansteigendes Zahlentripel gezogen wird (z.B. 18, 179, 657). Die Hypothese „ansteigende Zahlen“ wird durch das Zahlentripel 2, 4, 6 natürlich nicht widerlegt, erscheint aber doch unwahrscheinlich.

- 557 KLAYMAN und HA kritisieren WASON, weil die von ihm konstruierte Situation, in der die (nahe gelegte) Anfangshypothese enger ist als die gesuchte Hypothese, in Wirklichkeit kaum vorkomme. Meistens ist die getestete Hypothese zu weit, und die gesuchte Hypothese ist enger. In diesem Fall, wenn zudem positive Instanzen der Theorie seltener sind als negative Instanzen, ist die Suche nach positiven Instanzen der Theorie durchaus rational, weil sie häufiger zur Falsifizierung der Hypothese führt.⁸⁷⁹ Dass die positiven Instanzen einer Theorie seltener sind als die negativen, ist ebenfalls eine realistische Annahme: man sucht nach den Persönlichkeitsmerkmalen, die zur Straffälligkeit führen, oder nach Gründen, die zu ausserordentlichem geschäftlichem Erfolg führen, nicht nach Merkmalen oder Gründen, die zur Einhaltung der Gesetze oder durchschnittlichem Erfolg führen. Unter den Fällen, die von der Hypothese nicht vorausgesagt werden, nach falsifizierenden Instanzen zu suchen, gleicht hier der sprichwörtlichen Suche nach der Nadel im Heuhaufen.⁸⁸⁰ KLAYMAN und HA schlagen daher vor, die Strategie der Versuchspersonen in WASONs Zahlentripel-Aufgabe besser „positive Teststrategie“ statt „Bestätigungsstrategie“ zu nennen.⁸⁸¹

- 558 Das Zahlentripel-Experiment ist nicht geeignet, eine Bestätigungstendenz nachzuweisen, weil die Versuchspersonen nicht wissen können, ob die nachgefragte Information (die Antwort des Versuchsleiters) bestätigend oder widerlegend sein wird. Experimente, bei denen die Versuchspersonen von vorneherein wissen, ob die nachgefragte Information ihre Hypothese stützt oder nicht, leiden nicht an diesem methodischen Defizit und sind daher geeignet, eine Bestätigungstendenz bei der Informationssuche nachzuweisen. Zahlreiche Studien weisen eine solche verzerrte Informationssuche sowohl bei Individuen wie bei Gruppen nach.⁸⁸² Die bestätigende Tendenz ist umso stärker, je mehr die Versuchsperson an der zu überprüfenden Hypothese „hängt“. Dies wird – z.B. durch JONAS et al. – dissonanztheoretisch erklärt. Vereinfacht gesagt entstehen kognitive Dissonanzen dann, wenn zwei Wahrnehmungen eines Menschen mit einer subjektiven Hypothese über deren Beziehung im Widerspruch stehen. Wer an eine Theorie glaubt, hofft, dass die Fakten die Theorie stützen. Tun sie dies nicht, entsteht eine Dissonanz zwischen Theorie und Fakten, die aufgelöst werden muss. Um diese Dissonanz zu vermeiden, sucht man unbewusst gar nicht erst nach Informationen, die im Widerspruch zur Hypothese stehen könnten.

⁸⁷⁹ KLAYMAN/HA, FN 862, 217.

⁸⁸⁰ Ist aber in einem zweiten Schritt, wenn die Hypothese aufgrund der positiven Teststrategie nicht falsifiziert wurde, dennoch notwendig, da ansonsten der Likelihood-Quotient nicht gebildet werden kann.

⁸⁸¹ KLAYMAN/HA, FN 862, 212

⁸⁸² Nachweise bei EVA JONAS/STEFAN SCHULZ-HARDT/DIETER FREY, Konfirmatorische Informationssuche bei simultaner vs. sequentieller Informationsvorgabe, Zeitschrift für Experimentelle Psychologie 2001, 239-247.

- 559 Bei sequentieller Informationsvorgabe – d.h. die Versuchspersonen erhalten eine Information und sagen dann, welche weitere Information sie erhalten möchten – ist der *confirmation bias* stärker als bei gleichzeitiger Vorgabe der Informationen (d.h. die Versuchspersonen wählen aus den zur Verfügung stehenden Informationen diejenigen aus, die sie erfahren möchten).⁸⁸³ Gruppen, deren Mitglieder nicht alle die gleiche Präferenz teilen, suchen weniger ausgeprägt nach bestätigender Information. Die Heterogenität der Gruppe als solche, z.B. ob die Gruppe aus unterschiedlichen Persönlichkeiten oder Personen beider Geschlechts zusammengesetzt ist, hat hingegen keinen Einfluss auf den *confirmation bias*.⁸⁸⁴
- 560 CLOYD und SPILKER haben die bestätigende Informationssuche im juristischen Kontext untersucht.⁸⁸⁵ Ihre Versuchspersonen waren 71 Steuerberater der (damals noch) Big Five Buchprüfungsfirmen, die im Schnitt rund 2,5 Jahre Berufserfahrung hatten. Allen wurde der gleiche Sachverhalt vorgelegt, gemäss dem ein fiktiver Klient ein Grundstück verkauft hatte. Der Hälfte der Versuchspersonen wurde gesagt, dass der Klient beim Verkauf einen Gewinn von einer halben Million Dollar gemacht, der anderen Hälfte, dass er einen Verlust in gleicher Höhe erlitten habe. Der Klient, der einen Gewinn gemacht hat, möchte als „Investor“ im Sinne des anwendbaren Steuerrechts behandelt werden, weil dann der Gewinn als Kapitalgewinn steuerfrei ist. Umgekehrt möchte der Klient, der einen Verlust erlitten hat, als „Händler“ qualifiziert werden, weil er dann den geschäftsmässigen Verlust vom Einkommen absetzen kann. Der Sachverhalt liess beide Qualifikationen zu.⁸⁸⁶
- 561 Die Versuchspersonen hatten eine halbe Stunde Zeit, eine Datenbank nach einschlägigen Urteilen zu durchsuchen und die Urteile abzuspeichern, die sie in einem Memorandum zur Frage zitieren würden. Die speziell erstellte Datenbank enthielt gleich viele Urteile, in denen ein Gericht erkannt hatte, dass ein Steuerpflichtiger als Investor zu behandeln ist, wie Urteile, in denen erkannt wurde, dass er als Händler zu behandeln ist.⁸⁸⁷
- 562 Die Versuchspersonen in der „Investor“-Gruppe studierten während 16,6 Minuten Urteile, in denen ein Steuerpflichtiger als Investor qualifiziert wurde, aber nur während 6,2 Minuten Urteile, in denen entschieden wurde, dass der Steuerpflichtige ein Händler ist. Sie speicherten im Schnitt 4,5 „Investor“-Fälle gegenüber nur 1,6 „Händler“-Fälle ab. Die Gruppe „Händler“ andererseits betrachtete während 12,3 Minuten „Händler“-Fälle und nur während 7,6 Minuten „Investor“-Fälle. Die Versuchspersonen in dieser Gruppe speicherten im Schnitt 3,9 „Händler“- und 1,6 „Investor“-Fälle ab.⁸⁸⁸ Die Versuchspersonen wurden anschliessend gefragt, wie gross die Chance sei, dass ein Gericht die vom Klienten bevorzugte Sicht einnehmen würde. Die Gruppe „Händler“ sah eine Chance von 49 %, dass der

⁸⁸³ JONAS/SCHULZ-HARDT/FREY, FN 882, 244.

⁸⁸⁴ STEFAN SCHULZ-HARDT/DIETER FREY/KRISTIANE FAGO/GÜLER KICI, Selektive Informationssuche und Gruppenheterogenität: Der Einfluss verschiedener Formen der Gruppenheterogenität auf Selbstbestätigungsprozesse bei Entscheidungen, Gruppendynamik 1999, 161-174.

⁸⁸⁵ C. BRYAN CLOYD/BRIAN C. SPILKER, The Influence of Client Preferences on Tax Professionals' Search for Judicial Precedents, Subsequent Judgments and Recommendations, The Accounting Review 1999, 299-322.

⁸⁸⁶ CLOYD/SPILKER, FN 885, 307.

⁸⁸⁷ CLOYD/SPILKER, FN 885, 307.

⁸⁸⁸ CLOYD/SPILKER, FN 885, 309.

Klient als Händler qualifiziert würde, während die Gruppe „Investor“ eine Chance von 66 % sah, dass der Klient als Investor eingestuft würde. Die tatsächliche Wahrscheinlichkeit sollte selbstverständlich nicht von der Präferenz des Klienten abhängen. Weniger als 11 % der Versuchspersonen gaben eine Wahrscheinlichkeit zwischen 33 % und 66 % an, die bei einem ausgewogenen Sachverhalt und neutraler Beurteilung zu erwarten wäre.⁸⁸⁹ Eine Pfadanalyse zeigt, dass die unterschiedliche Beurteilung in erster Linie auf die verzerrte Informationssuche zurückzuführen ist.⁸⁹⁰

- 563 In einer weiteren Studie verwendeten CLOYD und SPILKER einen Sachverhalt, der eigentlich nur den Schluss zulies, dass der Klient als Händler einzustufen ist (vier neutrale Experten beurteilten unabhängig die Wahrscheinlichkeit, dass der Klient in diesem Fall von einem Gericht als „Investor“ behandelt würde, als zwischen 10 und 20 %). Die Experten waren der Meinung, dass eine Steuerabrechnung als „Investor“ in diesem Fall sogar zu Sanktionen führen würde. Allen Versuchspersonen der zweiten Studie wurde gesagt, dass der Klient eine Behandlung als Investor vorziehen würde, weil er einen Gewinn gemacht hatte. 46 % der Versuchspersonen empfahlen, dass der Klient als Investor abrechnen sollte. Diese Versuchspersonen verbrachten signifikant mehr Zeit als die Versuchspersonen, die die Abrechnung als Händler empfahlen, mit der Suche nach Urteilen, in denen Gerichte erkannt hatten, dass der Steuerpflichtige als Investor zu betrachten ist und speicherten signifikant mehr „Investor“-Urteile ab als die 54 % der Versuchspersonen, die zur richtigen Entscheidung gelangten.⁸⁹¹ Aufgrund der Korrelation zwischen Suchverhalten und Empfehlung darf geschlossen werden, dass das verzerrte Suchverhalten die (allzu aggressive) Empfehlung verursachte.⁸⁹²
- 564 Für Juristen möglicherweise tröstlich ist, dass eine weitere Untersuchung von CLOYD/SPILKER zeigt, dass Rechtsstudierende bei der Beurteilung eines ambivalenten Sachverhalts weniger als Studierende eines Buchhaltungslehrganges (*Master of Accounting*) zur konfirmatorischen Informationssuche neigen. Offenbar scheint die Ausbildung – zumindest an amerikanischen *Law Schools*, die nicht in jeder Hinsicht vergleichbar ist mit der Ausbildung an einer schweizerischen Rechtsfakultät – geeignet zu sein, gewissen „Wahrnehmungsstörungen“ entgegenzuwirken.⁸⁹³
- 565 KASSIN und Kollegen können zeigen, dass Verhörende, die von der Schuld des Verhörten überzeugt sind, aggressivere Einvernahmetechniken verwenden und mehr Druck auf den Verhörten ausüben; speziell wenn dieser (entgegen ihrer Vorstellung) unschuldig ist.⁸⁹⁴

⁸⁸⁹ CLOYD/SPILKER, FN 885, 310.

⁸⁹⁰ CLOYD/SPILKER, FN 885, 313.

⁸⁹¹ CLOYD/SPILKER, FN 885, 316.

⁸⁹² Dieser Schluss von Korrelation auf Kausation ist hier zulässig, weil in dem Laborexperiment die übrigen Variablen kontrolliert werden konnten.

⁸⁹³ C. BRYAN CLOYD/BRIAN C. SPILKER, Confirmation Bias in Tax Information Search: A Comparison of Law Students and Accounting Students, *Journal of the American Taxation Association* 2000, 60-71.

⁸⁹⁴ SAUL M. KASSIN/CHRISTINE C. GOLDSTEIN/KENNETH SAVITSKY, Behavioral Confirmation in the Interrogation Room: On the Dangers of Presuming Guilt, *Law and Human Behavior* 2003, 187-203.

Dies führt dazu, dass neutrale Beobachter das Verhalten der Verhörten als defensiver einstufen und die Verhörten als geringfügig schuldiger betrachten.⁸⁹⁵

II. Informationsbewertung

A. Ambivalente Informationen werden als Bestätigung interpretiert

- 566 Nicht jedes Beweismittel spricht eindeutig für oder gegen eine Hypothese. Gerade im Recht sind Informationen häufig, die so oder anders interpretiert werden können. Siehe nur das folgende Gedankenexperiment von CURT GOETZ:⁸⁹⁶

Sie gehen nachts spazieren und finden einen Mann in einer Blutlache liegen. Sie knien nieder ... ziehen ein Messer aus der Brust ... In diesem Augenblick hören sie Schritte, und plötzlich kommt ihnen ihre Situation zu Bewusstsein: ... der Tote kann nicht mehr für sie zeugen – und sie rennen davon, wie von Furien gejagt. Aber man fängt sie.

- 567 CURT GOETZ wünscht dem Betroffenen, dass er auf psychologisch geschulte Richter trifft. Man könnte auch sagen, auf Richter, die sich bewusst sind, dass Menschen die Neigung haben, ambivalente Informationen im Sinne der von ihnen als wahrscheinlich erachteten Hypothese zu interpretieren.⁸⁹⁷ Es ist offensichtlich, dass die Informationen in GOETZ' Gedankenexperiment ganz unterschiedlich bewertet werden können, je nachdem, ob man an die Schuld oder Unschuld des Gefassten glaubt.
- 568 Eines der bekanntesten Experimente zur Interpretation von Fakten im Lichte der eigenen Theorie stammt von LORD, LEPPER und ROSS.⁸⁹⁸ LORD und Kollegen wählten Versuchspersonen aus, die entweder Befürworter oder Gegner der Todesstrafe waren. Die Versuchspersonen erhielten zwei Studien zur Lektüre, die scheinbar empirische Belege für oder gegen die Todesstrafe lieferten. Sowohl Befürworter wie Gegner der Todesstrafe bewerteten die Studie, die ihre Position scheinbar stützte, als überzeugender und besser durchgeführt als die Studie, die gegen ihre Position sprach. Zweitens führte die Lektüre beider Studien dazu, dass die Überzeugungen der beiden Gruppen noch mehr polarisiert wurden; d.h. sowohl Befürworter als auch Gegner waren nach Kenntnisnahme aller Beweismittel – auch derjenigen, die ihrer Position widersprachen – von ihrer Position überzeugter als zuvor. Diese Polarisation entstand dadurch, dass die Überzeugung der Versuchspersonen jedes Mal erheblich zunahm, wenn sie ein Argument *für* ihre Position lasen, aber nur unwesentlich abnahm, wenn sie ein Argument gegen die Position lasen. KOEHLER hat in einer ähnlichen Studie Anhänger und Gegner parapsychologischer Phänomene gebeten, die Qualität von wissenschaftlichen Untersuchungen solcher Phänomene zu be-

⁸⁹⁵ KASSIN/GOLDSTEIN/SAVITSKY, FN 894, 198.

⁸⁹⁶ Zitiert nach BENDER/NACK, FN 605, X.

⁸⁹⁷ Zu dieser Tendenz LEE ROSS/CRAIG A. ANDERSON, Shortcomings in the Attribution Process, in: TVERSKY/KAHNEMAN/SLOVIC (Hrsg.), 129-152, 144 ff.; KLAYMAN, FN 855, 394.

⁸⁹⁸ CHARLES G. LORD/LEE ROSS/MARK R. LEPPER, Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence, Journal of Personality and Social Psychology 1979, 2098-2109.

werten. Beide Gruppen bewerteten Studien, die ihre Position stützten, als qualitativ besser.⁸⁹⁹

- 569 Zwei Drittel der Versuchspersonen von KOEHLER gaben an, dass ihre Bewertung von ihrer Einstellung zu den untersuchten Phänomenen nicht beeinflusst werde; 85 % meinten, dass ihre Bewertung nicht von der Einstellung beeinflusst werden sollte.⁹⁰⁰ KOEHLER weist (wie bereits LORD et al.) darauf hin, dass beide Aussagen falsch sind.⁹⁰¹ Es ist normativ richtig, die Resultate einer Studie mit einer gesunden Portion Skepsis zu betrachten, wenn sie einem theoretisch gut fundierten, durch zahlreiche unabhängige Experimente bestätigten (richtiger: nicht falsifizierten) Modell widersprechen. Wenn ein unbekannter Wissenschaftler berichtet, dass Schweine fliegen können, ist die Wahrscheinlichkeit, dass seine Beobachtung falsch ist, grösser, als dass alle Zoologen der Menschheitsgeschichte falsch lagen. Problematisch wird die Skepsis dann, wenn der Urteilende zwischen zwei Theorien entscheiden muss, die beide durch die vorliegenden (ambivalenten) Daten gestützt werden können, wenn man die Daten nur „richtig“ liest und keine der beiden Theorien *a priori* eine grössere Überzeugungskraft besitzt. Dies ist aber genau die Situation, in der sich der Richter oft befindet, wenn er über die Stichhaltigkeit der Theorien von Anklage und Verteidigung, von Kläger und Beklagtem, entscheiden muss. In diesem Fall fällt der Entscheid – in Grenzfällen – regelmässig auf die Seite, die der Richter, und sei es auch nur als „Arbeitshypothese“, ursprünglich als zutreffend erachtet.⁹⁰² Dies kann, siehe hinten, S. 188 ff., beispielsweise die These sein, die durch die zuerst vorgebrachten Beweismittel gestützt wird.

B. Bestätigende Informationen werden stärker gewichtet als widersprechende Informationen

- 570 Beweismittel, die für eine von der Versuchsperson gehaltene These sprechen, werden besser wahrgenommen, erinnert und stärker gewichtet als Beweismittel, die gegen die These sprechen.⁹⁰³ Nach GADENNE und OSWALD werden hypothesenbestätigende Informationen zu stark gewichtet, während hypothesekonträre Informationen erst als wichtig wahrgenommen werden, wenn sie der Hypothese deutlich widersprechen. Auch bei stärkster konträrer Information bleibt die anfängliche Hypothese immer noch die stärkste Hypothese; erst der deutliche Hinweis auf eine alternative Erklärung vermag den Effekt (weit-

⁸⁹⁹ JONATHAN J. KOEHLER, The Influence of Prior Beliefs on Scientific Judgements of Evidence Quality, *Organizational Behavior and Human Decision Making Processes* 1993, 28-55; siehe auch TOM PYSZCZYNSKI/JEFF GREENBERG, Toward an Integration of Cognitive and Motivational Perspectives on Social Inference: A Biased Hypothesis Testing Model, in: LEONARD BERKOWITZ (Hrsg.), *Advances in Experimental Social Psychology*, Vol. 30, San Diego etc. 1987, 297-340, 329.

⁹⁰⁰ KOEHLER, FN 899, 45 f.

⁹⁰¹ KOEHLER, FN 899, 30 f.; LORD/ROSS/LEPPER, FN 898, 2106.

⁹⁰² DEREK J. KOEHLER, Explanation, Imagination, and Confidence in Judgment, *Psychological Bulletin* 1991, 499-519.

⁹⁰³ NICKERSON, FN 855, 178; KLAYMAN, FN 855, 395; RICHARD NISBETT/LEE ROSS, *Human Inference: Strategies and Shortcomings of Social Judgment*, Englewood Cliffs 1980, 181 f.; MARGRIT E. OSWALD, Hypothesentesten: Suche und Verarbeitung hypothesenkonformer und hypothesekonträrer Informationen, in: HELL/FIEDLER/GIGERENZER (Hrsg.), FN 11, 189-212, 197; jeweils mit zahlreichen Hinweisen.

gehend) zum Verschwinden zu bringen.⁹⁰⁴ Gemäss der bereits vorne, S. 155, erwähnten Studie von SCHUM und MARTIN neigen Menschen dazu, bei der globalen Bewertung der Beweismittel konträre Beweismittel schlicht zu ignorieren oder als hypothesenbestätigend zu interpretieren.⁹⁰⁵ Man hat die Tendenz, sich darauf zu konzentrieren, wie gut die Beweismittel eine Hypothese stützen, und vergisst dabei zu überlegen, wie gut die gleichen Beweismittel auch mit einer alternativen Hypothese in Einklang stehen.⁹⁰⁶

- 571 LINDA JOHNSON legte 109 Steuerberatern vier Gerichtsurteile zu einer umstrittenen steuerlichen Abgrenzungsfrage vor. Der Sachverhalt blieb bei jedem Urteil immer derselbe, aber der Ausgang des Verfahrens wurde so manipuliert, dass in jeder Versuchsgruppe zwei andere Fälle zu einem für den Klienten günstigen Schluss kamen. Die Versuchspersonen wurden gebeten, die Relevanz der Urteile – die in erster Linie davon abhängt, ob die Fälle wegen des übereinstimmenden Sachverhalts gute Präjudizien sind – für den von ihnen zu beurteilenden Fall zu bewerten. Urteile, die zu einem für den Klienten positiven Schluss kamen, werden als relevanter beurteilt; diese Tendenz beeinflusste wiederum die Beurteilung der Prozesschancen durch die Steuerberater.⁹⁰⁷
- 572 Eine Studie von BABCOCK et al. belegt ebenfalls den Einfluss, den die Perspektive auf die Beurteilung der Prozesschancen haben kann. Wie vorne, S. 91, dargelegt, hängt es nach dem ökonomischen Modell von Vergleichsverhandlungen in erster Linie von der Einschätzung des Prozessausgangs durch die Parteien ab, ob die Parteien sich aussergerichtlich einigen. Die korrekte Voraussage des richterlichen Urteils hängt wiederum von den Fähigkeiten der Parteien ab, die Stärken und Schwächen ihres Falles unvoreingenommen zu beurteilen. Wenn bereits die Wahrnehmung des Falles verzerrt ist, werden auch die Urteilserwartungen systematisch verzerrt sein, und es kommt seltener zum Vergleich. BABCOCK et al. haben genau diese verzerrte Einschätzung der eigenen Prozesschancen nachgewiesen.⁹⁰⁸ Wenn alle ihre Versuchspersonen *zuerst* die Unterlagen zum Streitfall studierten und *dann* zufällig in die Gruppen „Beklagte“ oder „Kläger“ eingeteilt wurden, einigten sich 94 % der Paare auf einen Vergleich. Wurden die Versuchspersonen *vor* der Lektüre der Unterlagen in die Gruppen „Beklagte“ und Kläger eingeteilt, so einigten sich nur 72 % der Verhandelnden; ein signifikanter Unterschied.⁹⁰⁹ Die Differenz zwischen dem vom Kläger und vom Beklagten erwartetem Urteil war im ersten Fall im Schnitt nur \$ 6'936, im zweiten Fall aber \$ 18'555. Die Lektüre des Sachverhalts aus der Sicht einer

⁹⁰⁴ VOLKER GADENNE/MARGRIT OSWALD, Entstehung und Veränderung von Bestätigungstendenzen beim Testen von Hypothesen, Zeitschrift für experimentelle und angewandte Psychologie 1986, 360-374.

⁹⁰⁵ SCHUM/MARTIN, FN 770, 164.

⁹⁰⁶ DALE GRIFFIN/AMOS TVERSKY, The Weighing of Evidence and the Determination of Confidence, Cognitive Psychology 1992, 411-435; Nachdruck in: GILOVICH/GRIFFIN/KAHNEMAN (Hrsg.), 230-249, 238.

⁹⁰⁷ LINDA M. JOHNSON, An Empirical Investigation of the Effects of Advocacy on Preparers' Evaluations of Judicial Evidence, Journal of the American Taxation Association 1993, 1-22.

⁹⁰⁸ LINDA BABCOCK/GEORGE LOEWENSTEIN/SAMUEL ISSACHAROFF/COLIN CAMERER, Biased Judgments of Fairness in Bargaining, American Economic Review 1995, 1337-1342.

⁹⁰⁹ BABCOCK/LOEWENSTEIN/ISSACHAROFF/CAMERER, FN 908, 1340.

Partei hatte die Wahrnehmung der Versuchspersonen so verändert, dass ihre Einschätzung des Prozessausgangs systematisch zu ihren Gunsten verzerrt war.⁹¹⁰

III. Folgen der Bestätigungstendenzen

A. Festhalten selbst an einer diskreditierten Hypothese (*belief perseverance*)

- 573 Als *belief perseverance* wird das Phänomen bezeichnet, dass Leute selbst an einer völlig diskreditierten Theorie festhalten.⁹¹¹ In der ursprünglichen Studie von ROSS et al. wurden Versuchspersonen gebeten, echte von unechten Abschiedsbriefen von Selbstmördern zu unterscheiden. Der „Erfolg“ der Versuchspersonen war dabei rein zufällig und hing von der Gruppe ab, der sie zugeteilt worden waren. Auch nachdem sie darüber aufgeklärt worden waren, beurteilten die Versuchspersonen in der „Erfolgs“ Gruppe ihre Fähigkeit, ähnliche Aufgaben zu lösen, höher als die Versuchspersonen in der „Misserfolgs“ Gruppe. Zahlreiche weitere Studien bestätigten den ursprünglichen Befund von ROSS et al.⁹¹²

B. Einfluss der Reihenfolge der Beweismittel (*primacy effect*)

- 574 Die ursprüngliche Hypothese, die die Wahrnehmung und/oder Gewichtung von Beweismitteln beeinflusst, muss erst einmal generiert werden. Nicht in allen Fällen hat der Urteilende von vorneherein eine Meinung zum Thema; gerade im juristischen Bereich dürfte er häufig vorerst unvoreingenommen sein. Die Forschung zum *primacy effect* legt nahe, dass die Hypothese, die von dem oder den ersten Beweismitteln favorisiert wird, zur Arbeitshypothese wird und die Bewertung der weiteren Beweismittel beeinflusst.
- 575 Der *primacy effect* wurde erstmals von SOLOMON ASCH beschrieben. ASCH beschrieb der einen Hälfte seiner Versuchspersonen eine Person als intelligent-fleissig-impulsiv-kritisch-eigensinnig-neidisch, der anderen Hälfte als neidisch-eigensinnig-kritisch-impulsiv-fleissig-intelligent. Die Versuchspersonen in der ersten Gruppe beurteilten die beschriebene Person durchwegs positiver als diejenigen in der zweiten Gruppe.⁹¹³ ASCH selber interpretierte die Ergebnisse dahingehend, dass die Bedeutung der späteren Adjektive durch die zuerst wahrgenommenen Adjektive geformt wird; d.h. „eigensinnig“ wird positiv als „eigenständig“ kodiert, weil der betreffende Mensch intelligent ist und daher seinen Mitmenschen zu Recht häufig widerspricht. Diese Interpretation der Ergebnisse hat ihre Anhänger und Gegner; eine Gegenmeinung ist, dass die späteren Adjektive nicht anders codiert werden, sondern schlicht weniger Gewicht erhalten.⁹¹⁴ Die Annahme, dass der *primacy effect* durch die frühe Bildung einer Hypothese entsteht, ist mit beiden Auffassungen kom-

⁹¹⁰ Siehe auch GEORGE LOEWENSTEIN/SAMUEL ISSACHAROFF/COLIN CAMERER/LINDA BABCOCK, Self-Serving Assessments of Fairness and Pretrial Bargaining, *Journal of Legal Studies* 1993, 135-159.

⁹¹¹ LEE ROSS/MARK R. LEPPER/MICHAEL HUBBARD, Perseverance in Self Perception and Social Perception: Biased Attributional Processes in the Debriefing Paradigm, *Journal of Personality and Social Psychology* 1975, 880-892.

⁹¹² Nachweise bei NICKERSON, FN 855, 187; ROSS/NISBETT, FN 903, 175 ff.

⁹¹³ SOLOMON ASCH, Forming Impressions of Personality, *Journal of Abnormal and Social Psychology* 1946, 258-290, 270 f.

⁹¹⁴ Nachweise bei NISBETT/ROSS, FN 903, 173 f.

patibel.⁹¹⁵ Die ebenfalls nahe liegende Auffassung, dass die stärkere Beachtung zuerst dargebotener Beweismittel einfach darauf zurückzuführen ist, dass die Aufmerksamkeit der Versuchspersonen nachlässt, gilt als durch Experimente von JONES und GOETHALS widerlegt.⁹¹⁶

- 576 Studien, die einen *primacy effect* nachweisen, sind ausserordentlich zahlreich.⁹¹⁷ Manche Studien berichten aber auch von einem *recency effect*, d.h., die zuletzt dargebotenen Informationen beeinflussen das Urteil stärker. Dies erscheint paradox. Der *recency effect* scheint aber in erster Linie ein Gedächtnisphänomen zu sein; er tritt dann auf, wenn den Versuchspersonen relativ zahlreiche Informationen in schneller Folge dargeboten werden. In diesen Fällen erinnern sie sich nicht mehr an die ersten Informationen; diese können folglich auch das Urteil nicht mehr beeinflussen.⁹¹⁸ *Recency effects* treten zudem häufiger auf, wenn die Versuchspersonen nach jedem einzelnen Beweismittel ein Zwischenurteil abgeben, während *primacy effects* vorwiegend dann beobachtet werden, wenn die Beurteilung erst am Schluss, alle Informationen gesamthaft vorliegen, abgegeben wird.⁹¹⁹
- 577 Studien mit spezifisch juristischen Fragestellungen haben, wie andere Studien, sowohl *recency* wie *primacy effects* gefunden.⁹²⁰ PENNINGTON stellt im Gegensatz zu THIBAUT und WALKER einen klaren *primacy effect* fest und schreibt dies der grösseren Realitätsnähe seiner Studie zu.⁹²¹ Die neuste verfügbare Studie von COSTABILE/KLEIN findet wiederum *recency effects*, schreibt diese aber selber in erster Linie dem Gedächtnis der Geschworenen zu, die sich an Beweismittel, die erst spät eingeführt wird, im Zeitpunkt des Urteils besser erinnern.⁹²² Bei der Interpretation amerikanischer Studien zum Einfluss der Reihenfolge der Beweismittel auf Urteile von Geschworenen ist Vorsicht geboten, weil das amerikanische Prozessrecht sich grundlegend vom schweizerischen unterscheidet. In amerikanischen Geschworenenprozessen gilt das Unmittelbarkeitsprinzip und den Geschworenen ist es verboten, Notizen zu machen. Diese beiden Faktoren fördern natürlich *recency effects* stark; man kann sich vorstellen, wie schwierig es ist, sich nach einer Reihe von Zeugeneinvernahmen noch an die Aussagen des ersten Zeugen zu erinnern, wenn man keine Notizen machen darf.

⁹¹⁵ NISBETT/ROSS, FN 903, 174.

⁹¹⁶ NISBETT/ROSS, FN 903, 174 f.

⁹¹⁷ Nachweise bei NISBETT/ROSS, FN 903, 172 ff.; NICKERSON, FN 855, 187.

⁹¹⁸ NISBETT/ROSS, FN 903, 172.

⁹¹⁹ ROBIN M. HOGARTH/HILLEL J. EINHORN, Order Effects in Belief Updating: The Belief-Adjustment Model, *Cognitive Psychology* 1992, 1-55, 5.

⁹²⁰ DONALD C. PENNINGTON, Witnesses and Their Testimony: Effects of Ordering on Juror Verdicts, *Journal of Applied Social Psychology* 1982, 318-333, mit Nachweisen.

⁹²¹ PENNINGTON, FN 920, 330; JOHN THIBAUT/LAURENS WALKER, *Procedural Justice: A Psychological Analysis*, Hillsdale 1975, 61.

⁹²² KRISTI A. COSTABILE/STANLEY B. KLEIN, Finishing Strong: Recency Effects in Juror Judgments, *Basic and Applied Social Psychology* 2005, 47-58.

C. Zu grosse Sicherheit im Urteil (*overconfidence*)

- 578 Die ungenügende Berücksichtigung und Gewichtung widersprüchlicher Informationen kann dazu führen, dass sich der Entscheidende seiner Sache zu sicher ist (mehr zur Selbstüberschätzung generell hinten, S. 261 ff.).⁹²³ Frauen, die weniger zur Selbstüberschätzung neigen als Männer, bewerten hypothesenkonträre Informationen als wichtiger und empfinden (daher) die zu beurteilende Frage als schwieriger, was zu entsprechend geringerer empfundener Sicherheit bezüglich der Richtigkeit ihrer Antwort führt.⁹²⁴ Die Annahme, dass die einseitige Gewichtung der Argumente zu allzu grosser Gewissheit führt, dass das eigene Urteil richtig ist, wird durch Experimente gestützt, in denen die Versuchspersonen gezwungen werden, Gründe aufzulisten, die gegen die von ihnen getroffene Entscheidung sprechen.⁹²⁵ Wenn sie dies tun, stimmt ihre Einschätzung der Wahrscheinlichkeit, dass ihr Urteil richtig ist, eher mit der tatsächlichen Häufigkeit richtiger Urteile überein.⁹²⁶

IV. Eigene Studie zum Bestätigungsfehler

A. Frage zur Pseudodiagnostizität

- 579 Den Richterinnen und Richtern der Zivil- und Strafgerichte der Kanton beider Basel, Bern und Graubünden wurde folgender Sachverhalt vorgelegt:

Eine Fussgängerin wird von einem Auto angefahren und leicht verletzt. Der Fahrer des Autos begeht Fahrerflucht.

Die Fussgängerin sagt aus, sie sei entweder von einem Ford oder einem Opel angefahren worden. Sie kann nicht sagen, welche der beiden Marken es war, aber sie erinnert sich, dass es sich um einen schwarzen Kombi mit Schweizer Nummernschildern handelte.

- 580 Die Hälfte der Fragebogen („plausible Arbeitshypothese“) setzen sich wie folgt fort:

Sie wissen folgendes:

- A. Es gibt gleich viele Personenwagen der Marken Opel und Ford in der Schweiz.
- B. 60 % aller in der Schweiz zugelassener Personenwagen der Marke Opel sind Kombis.

Folgende Informationen sind ebenfalls erhältlich:

- C. Der prozentuale Anteil Kombis an der Gesamtheit in der Schweiz zugelassener Personenwagen der Marke Ford.
- D. Der prozentuale Anteil schwarzer Fahrzeuge an der Gesamtheit in der Schweiz zugelassener Personenwagen der Marke Opel.
- E. Der prozentuale Anteil schwarzer Fahrzeuge an der Gesamtheit in der Schweiz zugelassener Personenwagen der Marke Ford.

⁹²³ NICKERSON, FN 855, 189.

⁹²⁴ JANNE CHUNG/GARY S. MONROE, Gender Differences in Information Processing: An Empirical Test of the Hypothesis-Confirming Strategy in an Audit Context, *Accounting and Finance* 1998, 256-279.

⁹²⁵ ASHER KORIAT/SARAH LICHTENSTEIN/BARUCH FISCHHOFF, Reasons for Confidence, *Journal of Experimental Psychology: Human Learning and Memory* 1980, 107-118.

⁹²⁶ KORIAT/LICHTENSTEIN/FISCHHOFF, FN 925, 110.

581 Die andere Hälfte („unplausible Arbeitshypothese“) lautete folgendermassen:

Sie wissen folgendes:

A. Es gibt gleich viele Personenwagen der Marken Opel und Ford in der Schweiz.

B. 10 % aller in der Schweiz zugelassener Personenwagen der Marke Ford sind schwarz.

Folgende Informationen sind ebenfalls erhältlich:

C. Der prozentuale Anteil schwarzer Fahrzeuge an der Gesamtheit in der Schweiz zugelassener Personenwagen der Marke Opel.

D. Der prozentuale Anteil Kombis an der Gesamtheit in der Schweiz zugelassener Personenwagen der Marke Ford.

E. Der prozentuale Anteil Kombis an der Gesamtheit in der Schweiz zugelassener Personenwagen der Marke Opel.

582 In beiden Fällen wurden die Richterinnen und Richter gefragt, welche der zusätzlichen Informationen C, D oder E ihnen hilft entscheiden, ob die Fussgängerin einen Ford oder einen Opel gesehen hat, wenn keine anderen als die zur Auswahl stehenden Informationen erhältlich sind. In beiden Gruppen ist die Antwort C diagnostisch, die Antwort D pseudo-diagnostisch für die Arbeitshypothese und die Antwort E pseudodiagnostisch für die alternative Hypothese. Die Gruppen werden als „plausibel“ und „unplausibel“ bezeichnet, weil das bekannte Indiz „Art des Fahrzeuges“ in der ersten Gruppe bei mehr als 50 % der Instanzen der Kategorie „Opel“ vorliegt, während das bekannte Indiz „Farbe“ in der zweiten Gruppe nur bei 10 % der Instanzen der Kategorie „Ford“ vorliegt. Folglich erhält die *focal hypothesis* (hier als „Arbeitshypothese“ bezeichnet) durch das bekannte Indiz der ersten Gruppe mehr Unterstützung als durch das bekannte Indiz der zweiten Gruppe (in der ersten Gruppe $P(\text{Kombi} \mid \text{Opel}) = 0,6$; in der zweiten Gruppe $P(\text{Schwarz} \mid \text{Ford}) = 0,1$). Nach MYNATT, DOHERTY und DRAGAN müsste dies dazu führen, dass die alternative Hypothese eher in Betracht gezogen wird und daher der Anteil pseudodiagnostischer Antworten sinkt (womit auch gleich die Hypothese formuliert wäre).⁹²⁷

583 Systematisch lassen sich die bekannten und erhältlichen Informationen in einer Matrix gemäss Tabelle 14 darstellen. Die Gruppe „plausible Arbeitshypothese“ kennt die Information in der Zelle A ($P(\text{Kombi Opel}) = 0,6$). Diagnostisch ist die Information in Zelle B, während die Information in Zelle C weitere Informationen über die Arbeitshypothese („angenommen, es ist ein Opel“) liefert und gemäss üblichem Sprachgebrauch als pseudo-diagnostisch bezeichnet wird. Die Zelle D wird in traditionellen Pseudodiagnose-Studien selten gewählt; sie ist pseudodiagnostisch für die Alternativhypothese.

	Opel (H_1)	Ford (H_2)
% Kombis (I_1)	Zelle A: $P(I_1 \mid H_1)$ „Anteil Kombis an Opel“	Zelle B: $P(I_1 \mid H_2)$ „Anteil Kombis an Ford“
% schwarz (I_2)	Zelle C: $P(I_2 \mid H_1)$ „Anteil schwarz an Opel“	Zelle D: $P(I_2 \mid H_2)$ „Anteil schwarz an Ford“

Tabelle 14: Systematische Darstellung des Diagnose-Problems

⁹²⁷ MYNATT/DOHERTY/DRAGAN, FN 859, 766.

B. Resultate

584 190 der 234 Richterinnen und Richter, die den Fragebogen zumindest teilweise ausgefüllt zurückgeschickt hatten, beantworteten die Frage nach der Diagnostizität der Beweismittel. Die vergleichsweise hohe Ausfallquote von 18,8 % ist darauf zurückzuführen, dass relativ viele Richter meinten, keine der angebotenen Informationen würde es ihnen erlauben, die Frage, ob es sich um einen Opel oder einen Ford gehandelt habe, mit hinreichender Sicherheit zu entscheiden. Das ist zweifelsfrei richtig. Allerdings war die Frage, welche der Informationen *helfen* würde, zu entscheiden; und diese Frage kann meines Erachtens sinnvoll beantwortet werden. Sie mag zugegebenermassen wenig realistisch sein, weil man in einem tatsächlichen Gerichts- oder Untersuchungsverfahren den Aufwand, diese – auf jeden Fall nicht sehr diagnostischen – Informationen zu beschaffen, kaum treiben würde.

	Antworten			Total
	Diagnostisch (C)	PD für Arbeitsh. (D)	PD für Alternativh. (E)	
plausible Arbeitshypothese	34 (35 %)	44 (46 %)	18 (19 %)	96
unplausible Arbeitshypothese	38 (40 %)	26 (27 %)	30 (32 %)	94
	72	70	48	190

Tabelle 15: Resultate zur Frage nach dem diagnostischen Beweismittel (PD = pseudo-diagnostisch)

585 Die Resultate der Frage zur Diagnostizität der Beweismittel sind in Tabelle 15 dargestellt. Der Einfluss der Plausibilität der Arbeitshypothese auf die Antworten ist statistisch signifikant.⁹²⁸ Der Anteil der diagnostischen Antworten ist in beiden Gruppen ungefähr gleich gross (35 resp. 40 %). Der Trend – mehr diagnostische Antworten bei unplausibler Arbeitshypothese – zeigt in die erwartete Richtung. Ein erheblicher Unterschied zeigt sich bei den pseudodiagnostischen Antworten für die Arbeitshypothese; sie nehmen von 46 % bei plausibler Arbeitshypothese auf 27 % bei unplausibler Arbeitshypothese ab. Das Verhältnis von diagnostischen zu pseudodiagnostischen Antworten kehrt sich beim Wechsel von der plausiblen zur unplausiblen Arbeitshypothese um. Die Resultate von MYNATT, DOHERTY und DRAGAN wurden somit repliziert.⁹²⁹

586 Der Abnahme der Anzahl pseudodiagnostischer Antworten bei unplausibler Arbeitshypothese entspricht aber nicht, wie bei MYNATT et al., eine in etwa gleich starke Zunahme bei den diagnostischen Antworten. Stattdessen haben mehr Richterinnen und Richter in der Gruppe „unplausible Arbeitshypothese“ die pseudodiagnostische Information für die Alternativhypothese gewählt; nämlich 32 % gegenüber 19 %. Ein ähnliches Phänomen wurde von EVANS, VENN und FEENEY beobachtet, die ebenfalls einen stabilen Anteil an diagnostischen Antworten, aber eine Verschiebung zu Gunsten der sonst eher selten gewählten

⁹²⁸ $\chi^2 = 7,831$, $p = 0,02$.

⁹²⁹ MYNATT/DOHERTY/DRAGAN, FN 859, 774.

Antwort „E“ (bei ihnen „D“) feststellten, als sie die Plausibilität der *focal hypothesis* von stark zu schwach manipulierten.⁹³⁰

C. Diskussion

- 587 Die Resultate replizieren die Studie von MYNATT et al. und sprechen dafür, dass der Hauptgrund für die Wahl pseudodiagnostischer Informationen die Schwierigkeit ist, mehrere Alternativen gleichzeitig zu verfolgen. Stattdessen konzentriert sich der Mensch auf die Arbeitshypothese – wobei es bereits genügt, dass sich die verfügbaren Informationen auf diese Hypothese beziehen, damit eine Hypothese zur Arbeitshypothese wird – und sucht nach weiteren Informationen zu dieser Hypothese. Dabei vergisst man leicht, dass sich der abstrakte Beweiswert eines Indizes nur beurteilen lässt, wenn man weiss, wie oft es vorkommt, wenn die Arbeitshypothese nicht zutrifft.
- 588 Die beobachtete Verschiebung von Antworten „D“ zu Antworten „E“ bei der unplausiblen Hypothese wurde von MYNATT et al. nicht beobachtet, entspricht aber den Resultaten von EVANS et al. Diese erklären sie damit, dass die Versuchspersonen, die mit einer unplausiblen Arbeitshypothese konfrontiert sind, zu den Antworten „E“ wechseln, weil sie die Alternativhypothese zu ihrer neuen Arbeitshypothese machen. Die Antworten „E“ sind die pseudodiagnostischen Antworten für Leute, die die Alternativhypothese zu ihrer Arbeitshypothese gemacht haben.⁹³¹ So betrachtet verstärken diese Resultate die Überzeugungskraft der Theorie, dass pseudodiagnostische Informationssuche darauf zurückzuführen ist, dass man nicht mehr als eine Hypothese gleichzeitig im Arbeitsgedächtnis behalten kann.
- 589 Warum aber haben MYNATT et al. keine Verschiebung zugunsten von „E“ beobachtet? Eine spekulative, aber plausible, Erklärung ist, dass ihre Manipulation zur Diskreditierung der Plausibilität der Arbeitshypothese nicht ausreichend stark war, um einen Wechsel des Fokus zu bewirken. Bei MYNATT et al. betrug $P(I_1 | H_1)$ bei der plausiblen Hypothese 0,65 und bei der unplausiblen Hypothese 0,35 (also rund die Hälfte), während bei der vorliegenden Studie $P(I_1 | H_1)$ bei der plausiblen Hypothese 0,6 und bei der unplausiblen 0,1 – also sechs Mal weniger – betrug. Die Manipulation der Plausibilität der Arbeitshypothese war bei EVANS et al. m.E. ebenfalls stärker als bei MYNATT et al. Die Versuchspersonen bei EVANS et al. sollten aufgrund ganz ähnlicher Informationen, wie sie in der vorliegenden Studie verwendet wurden, herausfinden, welche Marke das Auto ihrer Schwester hat. Die Manipulation der Plausibilität der Arbeitshypothese geschah nicht über die Veränderung von $P(I_1 | H_1)$, sondern dadurch, dass den Versuchspersonen erklärt wurde, ihre Schwester lege grossen Wert auf ein sparsames Auto. Die Information ($I_1 | H_1$), die den Versuchspersonen bekannt war, war entweder, dass das Modell Y 14 Liter, 8 Liter oder 6 Liter Benzin pro 100 Kilometer verbrauche.⁹³² Da es eher unwahrscheinlich ist, dass jemand, der grossen Wert auf ein sparsames Auto legt, ein Auto besitzt, das 14 Liter Benzin pro 100 Kilometer verbraucht, ist die Arbeitshypothese im ersten Fall am unplausibelsten;

⁹³⁰ EVANS/VENN/FEENEY, FN 871, 36.

⁹³¹ EVANS/VENN/FEENEY, FN 871, 42.

⁹³² Im Original 20/25/50 (UK) mpg; EVANS/VENN/FEENEY, FN 871, 35.

die Versuchspersonen schlossen aus dieser Angabe wahrscheinlich, dass es sich beim Auto ihrer Schwester nicht um ein Modell Y handelt und konzentrierten sich auf das Modell X.

V. Was kann gegen Bestätigungstendenzen unternommen werden?

A. Kontrafaktisches Denken (*counterfactual reasoning*)

590 Als Strategie zur Vermeidung von Bestätigungstendenzen – und anderen kognitiven Täuschungen – hat sich das im Englischen als *counterfactual reasoning* bezeichnete Denken bewährt, das sich nur ungenügend mit „kontrafaktisches Denken“ ins Deutsche übersetzen lässt.⁹³³ Wenn man kontrafaktisch denkt, stellt man sich vor, welche Bedingungen hätten vorhanden sein müssen, damit ein Ereignis nicht eingetreten wäre.⁹³⁴ Kontrafaktische Gedanken rekonstruieren die Vergangenheit und werden meist durch Sätze, die mit „wenn (ich) doch nur (nicht)“ beginnen, ausgedrückt. Das Nachdenken darüber „was hätte sein können“ beeinflusst, wie wir über eine zukünftige Situation denken. Es schärft das Bewusstsein dafür, dass man mehr als eine Möglichkeit in Betracht ziehen sollte, ehe man eine Entscheidung fällt.⁹³⁵ Kontrafaktisches Denken erleichtert es, sich vorzustellen, welche Gründe *gegen* die derzeit favorisierte Hypothese sprechen.⁹³⁶ Menschen erzählen gerne und gut Geschichten; es fällt uns leicht, im Nachhinein eine Erklärung für eine Beobachtung zu finden, die plausibel und überzeugend klingt.⁹³⁷ Von der Plausibilität der Erklärung zur Überzeugung, dass die Erklärung richtig ist, ist es nur ein kleiner Schritt. Aktiv und bewusst Gründe zu suchen, die *gegen* die Erklärung/Hypothese sprechen, schärft den Blick für Informationen, die der ersten Erklärung widersprechen und gerne unter den Tisch gewischt werden. Es trägt dazu bei, die vorhandenen Informationen ausgewogener zu gewichten und wirkt daher einer der Ursachen von Bestätigungstendenzen entgegen.

B. Alternative Erklärungen generieren und testen

591 Alles deutet darauf hin, dass die Beachtung einer alternativen Hypothese den Bestätigungsfehler stark verringert oder gar verschwinden lässt.⁹³⁸ Erfolgreiche Versuchspersonen zeichnen sich dadurch aus, dass sie mehrere Hypothesen generieren und testen.⁹³⁹ Die Bevorzugung hypothesenbestätigender Informationen verschwindet, wenn die Versuchspersonen auf eine alternative Hypothese hingewiesen werden.⁹⁴⁰ Man sollte daher immer

⁹³³ LAURA J. KRAY/ADAM D. GALINSKY, The Debiasing Effect of Counterfactual Mind-Sets: Increasing the Search for Disconfirmatory Information in Group Decisions, *Organizational Behavior and Human Decision Processes* 2003, 69-81.

⁹³⁴ DANIEL KAHNEMAN/AMOS TVERSKY, The Simulation Heuristic, in: KAHNEMAN/SLOVIC/TVERSKY (Hrsg.), 201-208, 203.

⁹³⁵ KRAY/GALINSKY, FN 933, 70.

⁹³⁶ KORIAT/LICHTENSTEIN/FISCHHOFF, FN 925, 109.

⁹³⁷ KOEHLER, FN 902, 500 ff.

⁹³⁸ GADENNE/OSWALD, FN 904, 372; KLAYMAN, FN 855, 405; MYNATT/DOHERTY/DRAGAN, FN 859, 775; EVANS/VENN/FEENEY, FN 871, 42

⁹³⁹ KLAYMAN, FN 855, 405.

⁹⁴⁰ GADENNE/OSWALD, FN 904, 372.

versuchen, sich alternative Erklärungen für die beobachteten Fakten vorzustellen und zu testen. Auf keinen Fall sollte man sich frühzeitig gegenüber Dritten oder gar der Öffentlichkeit auf eine Hypothese festlegen; je grösser das *commitment* gegenüber einer Hypothese, desto unangenehmer ist es, die Hypothese wieder aufgeben zu müssen, und desto grösser die unbewusste Tendenz, alles zu tun, um dies zu vermeiden. Alternative Hypothesen können auch durch einen *advocatus diaboli* generiert werden, da es manchmal schwierig ist, sich alternative Hypothesen vorzustellen, wenn man sich bereits auf eine Erklärung eingeschossen hat.

§ 13 Nichtbeachten der Regression zum Mittelwert

- 592 In den frühen sechziger Jahren des letzten Jahrhunderts versuchte DANIEL KAHNEMAN, die Fluglehrer der israelischen Armee davon zu überzeugen, dass Lob das wirksamere Mittel als Tadel sei, die Leistungen der Schüler zu verbessern. Ein Fluglehrer liess ihn in unmissverständlichen Worten wissen, dass das nicht zutrifft.⁹⁴¹

Mit Verlaub, was Sie sagen, ist Unsinn. Oft habe ich Schüler überschwänglich gelobt, die ein ausgezeichnetes Manöver geflogen sind, und das nächste Mal sind sie fast immer schlechter geflogen. Und ich habe Schüler angeschrien, die ein schlechtes Manöver flogen, und in den meisten Fällen verbesserte sich ihre Leistung beim nächsten Mal. Erzählen sie mir nicht, dass Bestrafung nicht wirkt. Meine Erfahrung sagt das Gegenteil.

- 593 Ein weiteres Beispiel: Im März 2004 werden alle 200 Schüler eines Jahrgangs einer Primarschule einem standardisierten Lese-/Schreibtest unterzogen. Die 20 Schüler, die am schlechtesten abgeschnitten haben, erhalten ein Jahr lang speziellen Förderunterricht. Im März 2005 werden die gleichen 200 Schüler wieder getestet. Von den 20 Schülern, die beim ersten Test am schlechtesten abschnitten, sind beim zweiten Test nur noch vier bei den schlechtesten 20. Ein Erfolg für den Förderunterricht?
- 594 Die Ursache der beobachteten Veränderung ist in beiden Fällen ein statistisches Phänomen, das als Regression zum Mittelwert bezeichnet wird. Der Erfolg bei einem komplexen Flugmanöver und bei einer Prüfung hängt einerseits von den Fähigkeiten des Schülers ab. Er hängt aber auch von zufälligen (noch unerklärbaren) Faktoren ab – Probleme in der Beziehung oder eine leichte Erkältung können die Leistung verschlechtern, eine gute Tagesform kann sie verbessern. Die zufälligen Faktoren, die den Extremwert mitbestimmen haben, sind beim zweiten Test wahrscheinlich nicht bei den gleichen Schülern vorhanden. Da die zufälligen Faktoren sich bei den besten Schülern offensichtlich dahingehend ausgewirkt haben, dass ihre Leistung besser wurde und bei den schlechtesten Schülern, dass ihre Leistung schlechter wurde, werden sich die Leistungen sowohl der besten wie der schlechtesten Schüler bei der zweiten Messung Richtung Durchschnitt bewegen.⁹⁴²
- 595 Die Regression zum Mittelwert wurde erstmals von FRANCIS GALTON beschrieben.⁹⁴³ GALTON untersuchte den Zusammenhang zwischen der Körpergrösse von Kindern und ihrer Eltern. Er stellte fest, dass die Kinder sehr grosser Eltern⁹⁴⁴ in der Regel kleiner, die Kinder sehr kleiner Eltern aber grösser als ihre Eltern sind. Die Grösse der Kinder tendierte zum Mittelwert hin, was GALTON als „*regression towards mediocrity*“ bezeichnete. Regression zum Mittelwert kann z.B. ein Phänomen wie den „Fluch des *Sports Illustrated* Covers“ (*Sports Illustrated Jinx*) erklären: die Leistungen der meisten Sportler, die auf

⁹⁴¹ Zitiert nach HASTIE/DAWES, FN 399, 164; siehe auch DANIEL KAHNEMAN/AMOS TVERSKY, On the Psychology of Prediction, Psychological Review 1972, 237-251, 250.

⁹⁴² Wer einwendet, dass es Schüler gibt, die immer eine Sechs schreiben, hat Recht. Das liegt aber daran, dass die Prüfung nicht selektiv genug ist. Wäre die Skala nach oben offen, würden diese Schüler oft, aber nicht immer, mit einer Sieben oder Acht bewertet. Da ihre Fähigkeiten an einem „guten Tag“ für eine Acht reichen, genügt ihnen auch ein „schlechter Tag“, um eine Sechs zu schreiben.

⁹⁴³ FRANCIS GALTON, Regression Towards Mediocrity in Hereditary Stature, Journal of the Anthropological Institute 1886, 246-63.

⁹⁴⁴ GALTON verwendete als Prädiktor die Durchschnittsgrösse von Vater und Mutter.

dem Titelblatt von *Sports Illustrated* erscheinen, lassen danach nach. Der Grund dafür ist nicht in der schwarzen Magie von *Sports Illustrated* zu suchen, sondern darin, dass nur Sportler, die ausgezeichnete Leistungen – Extremwerte – erbracht haben, auf die Titelseite gehoben werden. Dass ihre Leistungen sich von diesen Extremwerten wieder mehr dem Durchschnitt annähern werden, ist ein statistisches Gesetz.⁹⁴⁵ Würde *Sports Illustrated* nur Sportler auf die Titelseite bringen, die sich an ihrem Karrieretiefpunkt befinden, könnte sich der Schluss aufdrängen, dass eine Titelgeschichte von *Sports Illustrated* die Leistungen verbessert...

- 596 Der lineare Zusammenhang zweier Variablen kann durch den so genannten Korrelationskoeffizienten r ausgedrückt werden.⁹⁴⁶ r kann Werte zwischen $+1$ und -1 annehmen. Ein Wert von $+1$ bedeutet, dass die zwei Variablen perfekt linear korrelieren, d.h. der Zunahme von Y um eine Standardabweichung entspricht eine Zunahme von X um eine Standardabweichung.⁹⁴⁷ Ein Wert von 0 bedeutet, dass kein (linearer) Zusammenhang zwischen den beiden Werten besteht. Der Korrelationskoeffizient erlaubt es, die erwartete Änderung von X bei bekannter Änderung von Y zu schätzen. Bei perfekter linearer Korrelation besteht dabei keine Unsicherheit; wenn Y bekannt ist, ist auch X bekannt. Nimmt der Durchmesser eines Kreises (Y) um 2 cm zu, so nimmt der Umfang (X) um $6,28\text{ cm}$ zu, der Zusammenhang zwischen den beiden Werten ist perfekt linear, für Regression zum Mittelwert bleibt kein Raum.
- 597 Ist der Korrelationskoeffizient 0 , d.h. besteht kein (linearer) Zusammenhang zwischen den Werten, so ist die Regression zum Mittelwert maximal. Die beste Schätzung, die man in Kenntnis des Wertes Y für den Wert X machen kann, ist der Durchschnitt von X , weil Y keinerlei Vorhersagen über den Wert von X erlaubt. Beispiel: wird eine faire Münze acht Mal geworfen, so erlaubt die Häufigkeit, mit der sie auf Kopf gelandet ist, keine Voraussage, wie häufig die Münze bei den nächsten acht Würfeln auf Kopf landen wird. Da die Münze fair ist, hat die Häufigkeit von Kopf in den ersten acht Würfeln keinerlei Zusammenhang mit der Häufigkeit von Kopf in den zweiten acht Würfeln.⁹⁴⁸ Die beste Schätzung für die Häufigkeit von Kopf in der zweiten Serie von acht Würfeln ist daher vier, der Durchschnitt von Kopf bei einer fairen Münze.
- 598 Liegt der Korrelationskoeffizient irgendwo zwischen 0 und 1 , z.B. bei $0,7$, so ist die beste Schätzung von X bei bekanntem Y , dass X um $r \cdot$ Standardabweichung vom Mittelwert entfernt ist.⁹⁴⁹ Der vorausgesagte Wert Y ist weniger weit (gemessen in Standardabweichungen) entfernt vom arithmetischen Mittel als der bekannte Wert X (die Standardabweichung ist ein Mass für die Entfernung eines Wertes vom Durchschnitt). Wenn man annimmt, dass zwischen den Ergebnissen des ersten Tests und des zweiten

⁹⁴⁵ PLOUS, FN 396, 118.

⁹⁴⁶ Die Berechnung von des Bravais-Pearson-Korrelationskoeffizienten wird in jedem Statistiklehrbuch dargestellt, siehe z.B. JANE FIELDING/NIGEL GILBERT, *Understanding Social Statistics*, London 2000, 165 ff. Für eine verständliche, nicht-mathematische Erklärung siehe WALTER KRÄMER, *Statistik verstehen*, 4. Aufl. München 2004, 183 ff.

⁹⁴⁷ Bei einem Wert von -1 entspricht die Zunahme des einen Wertes um einen bestimmten Betrag der Abnahme des anderen Wertes um den gleichen Betrag.

⁹⁴⁸ Beispiel von HASTIE/DAWES, FN 399, 164.

⁹⁴⁹ NISBETT/ROSS, FN 903, 151.

Tests im obigen Beispiel mit den Schülern ein Zusammenhang von $r = 0,7$ besteht – also eine hohe, aber keine perfekte, Korrelation – ist die korrekte, regressive Voraussage, dass der schlechteste Schüler des ersten Tests, der 2,58 Standardabweichungen vom Durchschnitt entfernt war,⁹⁵⁰ im zweiten Test „nur“ $0,7 \times 2,58 = 1,8$ Standardabweichungen vom Durchschnitt entfernt sein wird. Er wird damit weiterhin unter den schlechtesten 20 Schülern sein. Aber alle Schüler, die beim ersten Test besser als im siebten Rang klassiert waren, werden voraussichtlich beim zweiten Test nicht mehr zu den schlechtesten zwanzig Schülern gehören.⁹⁵¹ Von einem Erfolg des Förderunterrichts kann erst gesprochen werden, wenn diese Voraussage (die die Kenntnis der Korrelation der beiden Testwerte voraussetzt), übertroffen wird.

- 599 Wird der Einfluss der nicht perfekten Korrelation auf die Voraussage ausser Acht gelassen, spricht man von nicht regressiver Voraussage (*non regressive prediction*). Es lässt sich experimentell leicht nachweisen, dass Menschen der Regression zum Mittelwert zu wenig Beachtung schenken. Drei Gruppen von Versuchspersonen wurden beispielsweise gebeten, den Notenschnitt eines Studierenden aufgrund dreier unterschiedlicher Prädiktoren vorauszusagen: der ersten Gruppe wurde gesagt, dass der Student in der x-ten Perzentile seines Jahrgangs ist (z.B.: „seine Noten sind besser als die von 65 % seiner Mitstudierenden“), der zweiten Gruppe, dass der Student in der x-ten Perzentile bei einem Konzentrationstest abgeschnitten hat und der dritten Gruppe, dass der Student in der x-ten Perzentile bei einem Test abgeschnitten hat, der misst, wie humorvoll ein Mensch ist. Die Probanden wurden ausdrücklich darauf aufmerksam gemacht, dass das Abschneiden beim Konzentrationstest von äusseren Umständen abhängt und nicht sehr zuverlässig ist, sowie dass ein Zusammenhang zwischen Humor und Notenschnitt nicht erwiesen ist. Die Voraussagen der Versuchspersonen, die auf dem mentalen Konzentrationstest beruhten, waren nicht regressiver als die Voraussagen, die auf dem Klassenrang beruhten, obwohl der Konzentrationstest eine unzuverlässige Messmethode ist. Selbst die auf dem Humor-Test basierenden Voraussagen des Notenschnitts korrelierten mit 0,7 mit dem Resultat des Humor-Tests, obwohl der Humor-Test keine valablen Aussagen über den zu erwartenden Notenschnitt erlaubt.⁹⁵²
- 600 KAHNEMAN und TVERSKY argumentieren, dass regressive Voraussagen aus drei Gründen kontraintuitiv sind.⁹⁵³ Es scheint vernünftig, dass eine optimale Regel der Voraussage es zumindest erlauben, wenn nicht sogar garantieren, sollte, dass alle Fälle korrekt vorausgesagt werden. Die Regression zum Mittelwert schliesst dies aus: weil die Varianz der Voraussagen immer geringer ist als die Varianz der Prädiktoren, ist klar, dass die Extremwerte der abhängigen Variable nicht korrekt vorausgesagt werden. Die Regression zum Mittelwert garantiert einen systematischen Fehler (der aber minimiert ist). Zweitens scheint es einleuchtend, dass die Beziehung zwischen einem beobachteten Wert und dem aufgrund der Beobachtung vorausgesagten Wert umkehrbar ist. Wenn B aufgrund von A vorausge-

⁹⁵⁰ Unter der Annahme, dass die Testresultate normalverteilt waren.

⁹⁵¹ Der Schüler im 20. Rang ist -1,28 Standardabweichungen vom Durchschnitt entfernt. $-1,28 \cdot (1/0,7) = 1,83$, was der 3,36 % Perzentile, oder dem siebten Rang von unten ($0,0336 \cdot 200 = 6,72$), entspricht.

⁹⁵² KAHNEMAN/TVERSKY, FN 941, 247.

⁹⁵³ DANIEL KAHNEMAN/AMOS TVERSKY, On the Study of Statistical Intuitions, *Cognition* 1982, 123-141, 137 ff.

sagt wurde, sollte A vorausgesagt werden können, wenn B bekannt ist. Regressive Voraus-sagen sind aber nicht umkehrbar (genauso wenig wie bedingte Wahrscheinlichkeiten). Wenn die vorausgesagte Grösse des Sohnes 1,87 m ist, weil der Vater 1,95 m gross ist, so ist die vorausgesagte Grösse des Vaters, wenn bekannt ist, dass der Sohn 1,87 m ist, nicht 1,95 m – sondern liegt zwischen 1,87 m und 1,76 m, der Durchschnittsgrösse Schweizer Männer.

- 601 Drittens möchten Menschen für eine Wirkung eine Ursache haben. Wir mögen Geschichten, nicht Statistik. Wenn man weiss, dass jemand bei einer Prüfung sehr gut abgeschnitten hat, bedeutet dies, dass er bei einer Wiederholung des Tests wahrscheinlich schlechter abschneiden wird. Das gute Resultat bei der ersten Prüfung ist aber natürlich nicht der *Grund* für das schlechte abschneiden beim zweiten Mal. Wir suchen deshalb nach Gründen wie „es ist ihm in den Kopf gestiegen“, „sie hat sich keine Mühe mehr gegeben“ oder ähnlich. KAHNEMAN/TVERSKY meinen:⁹⁵⁴

Eine Wirkung ohne Ursache verletzt eine starke Intuition. Das Verständnis der Regression zum Mittelwert wird schwer behindert dadurch, dass jedes Mal, wenn eine Regression beobachtet wird, eine Ursache für den beobachteten Effekt gesucht wird.

- 602 Die Annahme von KAHNEMAN/TVERSKY, dass die Nichtbeachtung der Regression zum Mittelwert zu härteren Strafen und weniger Belohnungen führt, wurde erstmals 2001 von NOTZ und Kollegen empirisch getestet.⁹⁵⁵ Die Versuchspersonen konnten einen „Untergebenen“ belohnen, wenn seine Leistung einen bestimmten Schwellwert überschritt, indem sie ihm einen Bonus zahlten, oder bestrafen, wenn seine Leistung eine bestimmte Schwelle nicht erreichte. Die Belohnung oder Bestrafung hatte dabei keinerlei Einfluss auf die Leistung des Untergebenen, was die Versuchspersonen nicht wussten und auch nicht erkannten.⁹⁵⁶ NOTZ et al. stellten eine interessante, nicht vorausgesagte, Wechselwirkung zwischen dem Talent des Untergebenen und der Belohnung/Bestrafung fest: bei den angeblich talentierten Untergebenen blieb das Ausmass der Belohnung konstant, aber die Bestrafung nahm bei jedem Fehler zu. Bei den angeblich untalentierten Untergebenen nahm die Belohnung ab, aber die Bestrafung blieb gleich. Die Hypothese, dass die Bestrafung generell zu und die Belohnung generell abnehmen würde, konnte nicht bestätigt werden.⁹⁵⁷ NOTZ et al. bezweifeln deshalb, dass die Nichtbeachtung der Regression zum Mittelwert alleine genügt, die weit verbreitete Bevorzugung von Bestrafung als Mittel der Verhaltensänderung zu erklären.

I. Regression zum Mittelwert und harte Strafen

- 603 Die Nichtbeachtung der Regression zum Mittelwert ist heimtückisch, weil Massnahmen, die einen Zustand verbessern sollen, meist dann erfolgen, wenn sich der Zustand einem (negativen) Extremwert nähert. Der Fussballtrainer wird gefeuert, wenn die Mannschaft

⁹⁵⁴ KAHNEMAN/TVERSKY, FN 953, 138.

⁹⁵⁵ WILLIAM W. NOTZ/IRVIN BOSCHMANN/NEALIA S. BRUNING, Punishment Without Cause: Regression and the Effects of Leader Attribution Error, *Journal of Applied Social Psychology* 2001, 2401-2416.

⁹⁵⁶ NOTZ/BOSCHMANN/BRUNING, FN 955, 2408.

⁹⁵⁷ NOTZ/BOSCHMANN/BRUNING, FN 955, 2412.

nur noch verliert, der Manager, wenn die Gewinne eingebrochen sind, und der Drogensüchtige steht vor dem Richter, wenn es ihm wieder einmal ganz besonders dreckig geht. Nach diesen Extremwerten kann es fast nur noch aufwärts gehen, und es ist sehr schwierig zu sagen, ob der beobachtete Aufwärtstrend am neuen Coach der Fussballmannschaft, an der Bestrafung des Drogensüchtigen oder an der Regression zum Mittelwert liegt.

- 604 FITZMAURICE/PEASE waren die ersten, die darauf hinwiesen, dass die Nichtbeachtung der Regression zum Mittelwert Richter glauben lassen kann, dass ihre Strafen die beobachtete Verhaltensänderung bewirken.⁹⁵⁸ Die kriminologische Forschung hat bekanntlich keinen Zusammenhang zwischen Härte der Sanktion und Rückgang der Straftaten gefunden.⁹⁵⁹ Nach der bundesgerichtlichen Rechtsprechung dürfen generalpräventive Zwecke bei der Strafzumessung berücksichtigt werden, aber nur in dem Masse, als dem Verurteilten ein Schuldvorwurf gemacht werden kann (BGE 118 IV 342, 350). Wenn Richter die Regression zum Mittelwert nicht beachten, kann sich bei ihnen leicht der Eindruck halten, dass ihre Strafen eine bessernde Wirkung (sei es spezial- oder generalpräventiver Art) haben. Die Nichtbeachtung der Regression zum Mittelwert legt es daher nahe, dass Gerichte sich bei der Strafzumessung stärker von general- und spezialpräventiven Motiven leiten lassen, als in der Urteilsbegründung ausdrücklich erwähnt und durch die kriminologische Forschung angezeigt. Die Meinung, dass harte Strafen ein wirksames Mittel zur Verhaltenssteuerung sind, ist jedenfalls weit verbreitet. Die Beratungsstelle für Unfallverhütung begrüsst den weiter hinten besprochenen BGE 130 IV 58, der zwei Autofahrer, die sich auf öffentlichen Strassen ein Rennen geliefert und dabei einen tödlichen Unfall verursacht haben, wegen eventualvorsätzlicher Tötung zu 6,5 Jahren Zuchthaus verurteilte, ausdrücklich als „wichtigen Beitrag zur Erhöhung der Verkehrssicherheit!“ – Ausrufzeichen im Original.⁹⁶⁰ Einen solchen Beitrag kann das Urteil aber nur leisten, wenn es abschreckend wirkt, was aus verschiedenen, hinten S. 208 f. näher dargelegten Gründen zu bezweifeln ist.
- 605 Es ist kaum möglich, empirisch nachzuweisen, ob Richter bei der Strafzumessung von der Nichtbeachtung der Regression zum Mittelwert beeinflusst werden. Erstens ist es zwar nahe liegend, dass ein Rückgang von Straftaten nach einem Extremwert auf die Regression zum Mittelwert zurückzuführen ist, aber ob die Regression zum Mittelwert die einzige Ursache ist, ist damit natürlich noch nicht bewiesen. Es kann nicht ausgeschlossen werden, dass auch harte Strafen, die auf dem Höhepunkte der Kriminalitätswelle ausgesprochen wurden, zum Rückgang beigetragen haben. Die Wirkung von Strafe und Regression zum Mittelwert lässt sich ausserhalb streng kontrollierter Laborbedingungen kaum trennen. Zweitens begründen Gerichte in der Schweiz die Strafzumessung im konkreten Einzelfall meist nicht explizit mit generalpräventiven Überlegungen, sondern stellen das Verschulden

⁹⁵⁸ FITZMAURICE/PEASE, FN 185, 38.

⁹⁵⁹ HANS WIPRÄCHTIGER, in: MARCEL ALEXANDER NIGGLI/HANS WIPRÄCHTIGER (Hrsg.), Strafgesetzbuch I: Art. 1-110 StGB, Basel etc. 2002, Art. 63 N 40: „Was die negative Generalprävention („Abschreckung“) betrifft, bestehen empirisch keine hinreichenden Hinweise dafür, dass eine Strafsteigerung einen feststellbaren Abschreckungseffekt hätte“; noch deutlicher Art. 63, N 47: „Härtere Strafen tragen nicht zur Kriminalitätsverminderung bei“.

⁹⁶⁰ Beratungsstelle für Unfallverhütung, bfu aktuell 2004/5, Bern 2004, 13 (erhältlich unter www.bfu.ch/medien/bfu_aktuell/2004_05_d.pdf, besucht am 11. Februar 2005).

des Verurteilten ins Zentrum. Ob sie sich in Wahrheit doch (auch) von generalpräventiven Zwecken leiten liessen, lässt sich nur vermuten, nicht belegen. Es wurde meines Wissens denn auch noch nie versucht, die Wirkung der Nichtbeachtung der Regression zum Mittelwert auf die Strafzumessung empirisch zu erforschen. Die Resultate von NOTZ et al. lassen auch daran zweifeln, ob der Zusammenhang so klar und eindeutig ist, wie von KAHNEMAN und TVERSKY ursprünglich angenommen.

A. Das Problem der Raserunfälle

- 606 Die folgende Geschichte ist aus den oben genannten Gründen genau das: eine Geschichte, kein Beweis. Ich bin mir auch der Gefahr bewusst, dass Menschen die Wirklichkeit so sehen, dass sie ihrem Weltbild entspricht.⁹⁶¹ Der Leser möge daher selbst entscheiden, ob er die Geschichte für überzeugend hält. Zusammengefasst geht sie so: seit einigen Jahren wird in der Presse in einem unerhörten Ausmass über so genannte „Raserunfälle“ – Auto-unfälle mit meist schweren Folgen, die auf massiv übersetzte Geschwindigkeit zurückzuführen sind – berichtet, so dass der (durch die Verkehrsunfallsstatistik nicht bestätigte) Eindruck entsteht, dass die Anzahl solcher Unfälle massiv gestiegen ist. In diesem Klima des medialen Hypes bestraft das Bundesgericht zwei Autoraser, die einen tödlichen Unfall verursacht haben, trotz massivster dogmatischer Bedenken wegen eventualvorsätzlicher Tötung. Dieses Urteil lässt sich wohl nur dadurch erklären, dass das Bundesgericht sich von der hohen Strafe eine abschreckende Wirkung erhofft. Dass die „Welle“ der Raserunfälle – besser gesagt: die Welle der Berichterstattung darüber – wieder abflachen wird, ist klar. Dass die Regression zum Mittelwert viel mehr dazu beitragen wird als die (rechtsstaatlich bedenkliche) neue Rechtsprechung des Bundesgerichts, scheint mir auch nahe liegend.

1. Verkehrsunfallstatistik

- 607 Die Zahl der in der Schweiz durch Strassenverkehrsunfälle getöteten Personen hat seit 1970 ständig abgenommen. Dies zeigen die indexierten Zahlen (1990 = 100) der durch Verkehrsunfälle in der Schweiz getöteten, leicht und schwer verletzten Personen (Abbildung 12).⁹⁶² In absoluten Zahlen hat die Zahl der Getöteten von 1970 bis 2003 von 1'649 auf 549 und der Schwerverletzten von 18'314 auf 5'862 abgenommen, während die Zahl der Leichtverletzten von 17'667 auf 24'236 zugenommen hat. Gleichzeitig hat die Zahl der in der Schweiz zugelassenen Strassenfahrzeuge (Personenwagen, Personen- und Sachentransportfahrzeuge, Landwirtschafts- und Industriefahrzeuge und Motorräder) von 1'666'143 im Jahr 1970 auf 4'888'296 im Jahr 2003 zugenommen, wobei die Personenwagen den prozentual grössten Anteil am Gesamtbestand und an der Zunahme ausmachen (2003: 3'753'890 zugelassene Personenwagen; 77 % des Gesamtfahrzeugbestands). Insgesamt gesehen hat die Verkehrssicherheit in den letzten 30 Jahren in der Schweiz daher

⁹⁶¹ LINDA CHAPMAN/JOHN CHAPMAN, Genesis of Popular but Erroneous Diagnostic Observations, *Journal of Abnormal Psychology* 1967, 193-204.

⁹⁶² Quelle (immer wenn nicht anders angegeben): Bundesamt für Statistik, Strassenverkehrsunfälle in der Schweiz 2003, Neuchâtel 2004.

stetig und stark zugenommen; wäre die Anzahl der im Strassenverkehr getöteten Personen mit dem Fahrzeugbestand gestiegen, so würden heute jährlich 4'840 Menschen durch Strassenverkehrsunfälle getötet.

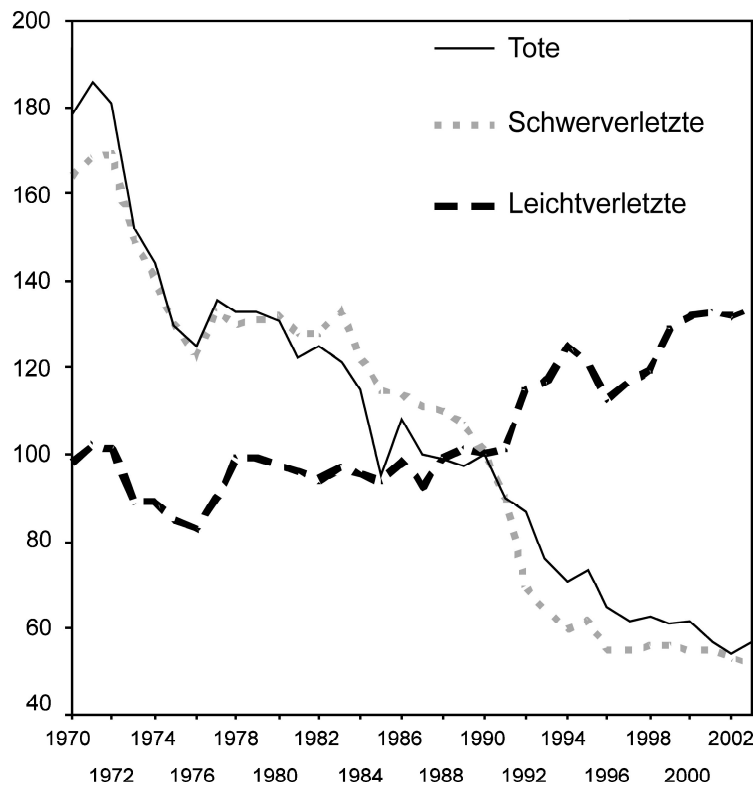


Abbildung 12: Unfallfolgen indexiert (1990 = 100)

608 Eine Analyse der Zahlen seit 1992 zeigt, dass sowohl die Zahl der Toten (1992: 834, 2003: 546 Getötete) als auch diejenige der Schwerverletzten weiterhin abgenommen hat (1992: 7'844, 2003: 5'862 Schwerverletzte). Von 2002 zu 2003 hat jedoch die Zahl der Getöteten von 513 auf 546 zugenommen (die Zahl der Schwerverletzten hat weiterhin abgenommen, von 5'931 auf 5'862). Die Zahlen des Bundesamtes für Statistik (BFS) für 2004 waren im Zeitpunkt der Redaktion dieser Arbeit noch nicht erhältlich. Die Statistik der Beratungsstelle für Unfallverhütung (bfu), die auf einer anderen Erhebungsmethode beruht, deren Ergebnisse für die Vorjahre aber identisch mit den Zahlen des BFS sind, zeigt einen Rückgang auf 509 Tote für das Jahr 2004.⁹⁶³ Betrachtet man die Ursachen der tödlichen Unfälle seit 1992, so sind zwei Ursachen für den Grossteil der Todesfälle verantwortlich: Geschwindigkeit und Zustand des Fahrers (Fahren in angetrunkenem Zustand, unter Drogen Einfluss, Übermüdung). Die Zahl der durch überhöhte Geschwindigkeit getöteten Personen sank bis 1997 stetig und stabilisierte sich dann. Von 2000 zu 2001 ist ein starker Rückgang sichtbar, der aber von 2001 bis 2003 wieder „aufgeholt“ wird. Aufgrund der relativ geringen Anzahl von Datenpunkten ist der Rückgang von 2000 auf 2001 möglicherweise zufällig. Er führt aber dazu, dass seit 2001 die Zahl der durch überhöhte Geschwindigkeit getö-

⁹⁶³ Beratungsstelle für Unfallverhütung, Medienmitteilung vom 11. Februar 2005; erhältlich unter www.bfu.ch/medien/infomedia/index.htm (besucht am 17. Februar 2005).

teten Personen ständig steigt (siehe Abbildung 13; bis zu drei Gründe pro Unfall werden erfasst, daher kann es zu Mehrfachzählungen eines Unfalls kommen).

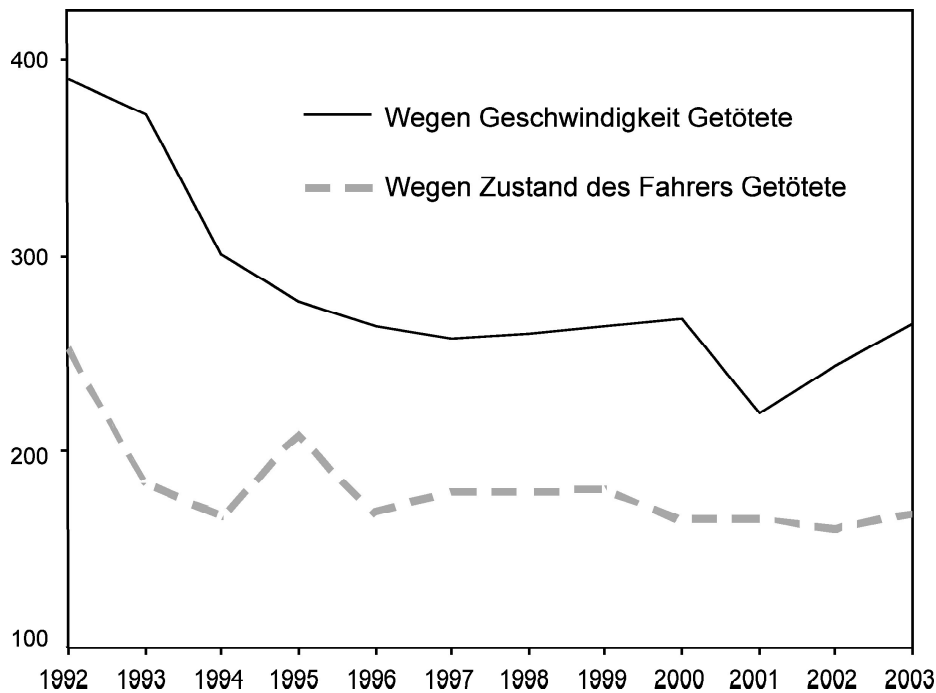


Abbildung 13: Todesursachen der im Strassenverkehr getöteten Personen seit 1992

2. Strafurteilsstatistik

609 Jeder dritte Mann in der Schweiz wird statistisch betrachtet einmal in seinem Leben wegen eines SVG-Delikts zu einer Strafe verurteilt, die zu einem Eintrag im zentralen Strafregister des Bundes führt. Nach Angaben des BFS ist die Anzahl der Verurteilungen gemessen am Fahrzeugbestand seit den 60-er Jahren stabil, aber die Härte der ausgesprochenen Strafen nimmt zu.⁹⁶⁴ Meine eigenen Kalkulationen, die auf die Daten seit 1984 beschränkt sind, zeigen, dass die Rate der Verurteilungen pro zugelassenes Fahrzeug von 1989 bis 1997 gesunken ist und seitdem wieder stark zunimmt (Abbildung 14).⁹⁶⁵

⁹⁶⁴ Bundesamt für Statistik, Medienmitteilung vom Februar 1999; erhältlich unter www.bfs.admin.ch/bfs/portal/de/index/news/medienmitteilungen.html (besucht am 24. März 2005).

⁹⁶⁵ Alle Zahlen für die Verurteilungen stammen aus der Strafurteilsstatistik 2004 des BFS; erhältlich unter www.bfs.admin.ch/bfs/portal/de/index/themen/rechtspflege/kriminalitaet__strafvollzug.html (besucht am 17. Februar 2005).

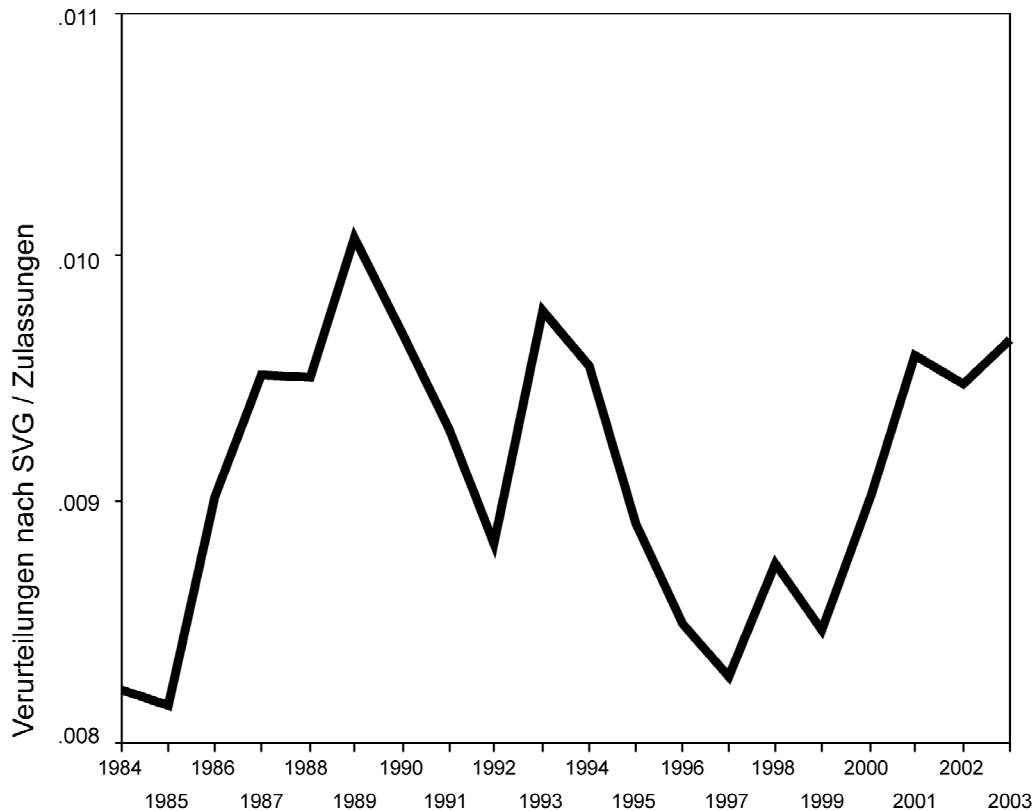


Abbildung 14: Anzahl Verurteilungen nach SVG pro zugelassenem Fahrzeug

- 610 In absoluten Zahlen haben die Verurteilungen wegen SVG-Delikten in den letzten 20 Jahren stetig zugenommen (Abbildung 15). Die bei Strassenverkehrsdelikten verhängten Sanktionen wurden zudem in den 90er-Jahren immer härter. Im Jahr 1990 wurden 36 % der Erstverurteilungen wegen Fahrens in angetrunkenem Zustand mit Busse und 64 % mit einer bedingten oder unbedingten Freiheitsstrafe geahndet. Zwischen 1990 und 1999 hat sich der Anteil der Verurteilungen zu Freiheitsstrafen von 64 % auf 73 % erhöht. Im gleichen Zeitraum stieg die durchschnittliche Dauer der unbedingten Freiheitsstrafen von 16 auf 22 Tage und jene der bedingten Freiheitsstrafen von 35 auf 51 Tage.⁹⁶⁶ Auch die Anzahl der Fahrausweisentzüge wegen Geschwindigkeitsübertretung, die bis 1997 stabil geblieben war, stieg von 1997 bis 2001 stark an (+62 %).⁹⁶⁷

⁹⁶⁶ Bundesamt für Statistik, Medienmitteilung vom 27. August 2002.

⁹⁶⁷ Bundesamt für Statistik, Medienmitteilung vom 27. August 2002.

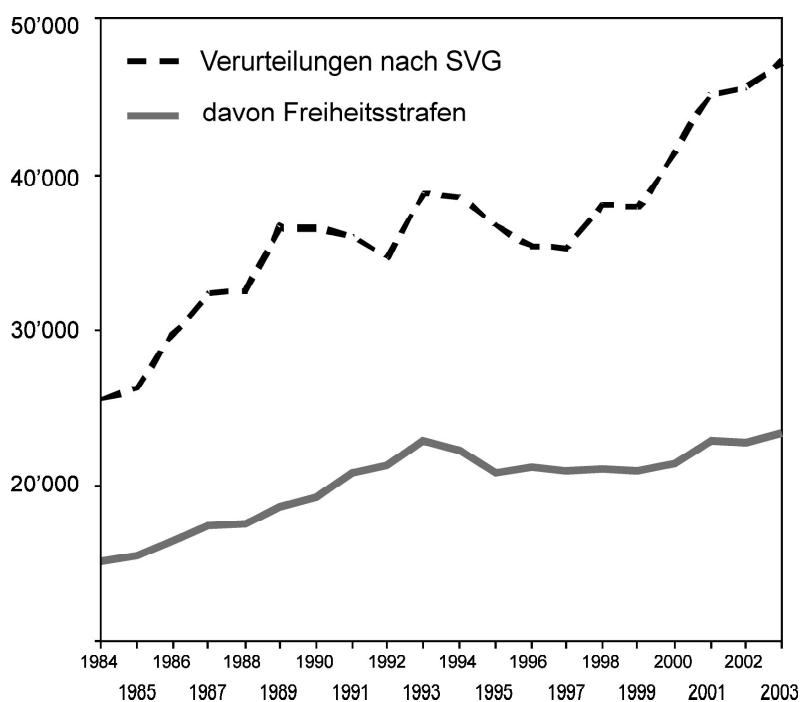


Abbildung 15: Verurteilungen nach SVG in absoluten Zahlen

611 Härtere Strafen haben keinerlei Einfluss auf die Rückfallwahrscheinlichkeit. Da die Sanktionen in der Schweiz von Kanton zu Kanton sehr unterschiedlich sind, lässt sich dies relativ einfach statistisch nachweisen. Der Bussenanteil an allen verhängten Sanktionen wegen Fahrens in angetrunkenem Zustand reicht von 97 % bis zu 4 %, wobei ausgeschlossen werden kann, dass es in einem Kanton nur gute Prognosen gibt (die für eine Busse sprechen) und in anderen Kantonen nur schlechte; vielmehr handelt es sich um ein Abbild der kantonalen Sanktionsstile.⁹⁶⁸ In Kantonen, die häufiger Freiheitsstrafen verhängen, kommt es weder zu mehr noch zu weniger Wiederverurteilungen als in den Kantonen, in denen bei Fahren in angetrunkenem Zustand hauptsächlich Bussen ausgesprochen werden.⁹⁶⁹ Ein Viertel der wegen eines Strassenverkehrsdelikts verurteilten Personen wird innerhalb von sieben Jahren erneut verurteilt. Diese auf den ersten Blick hohe Zahl relativiert sich stark, wenn man sich bewusst macht, dass rund ein Drittel der Männer mindestens ein Mal in ihrem Leben wegen eines Verkehrsdelikts verurteilt wird. Das Bundesamt für Statistik beurteilt die Wiederverurteilungsrate bei Strassenverkehrsdelikten denn auch als „relativ bescheiden“.⁹⁷⁰

3. Öffentliche Wahrnehmung der und Berichterstattung über „Raserunfälle“

612 Die Wahrnehmung der Raserunfälle als Problem hat in den letzten drei Jahren explosionsartig zugenommen, obwohl sie sich in der Verkehrsunfallstatistik bislang nicht widerspiegeln (wobei man ehrlicherweise anfügen muss, dass die Zahlen für 2004 noch nicht erhältlich

⁹⁶⁸ STEVE VAUCHER/RENA TE STORZ/SIMONE RÔNEZ, Strassenverkehrsdelinquenz und Rückfall, Neuchâtel 2000, 20.

⁹⁶⁹ A. a. O.

⁹⁷⁰ A. a. O., 25.

waren, als diese Zeilen geschrieben wurden). Um diese subjektive Empfindung halbwegs zu objektivieren, habe ich die Mediendatenbank *Swissdox*, die die 50 wichtigsten schweizerischen Tageszeitungen und Wochenzeitschriften erfasst,⁹⁷¹ nach dem Wort „Raserunfall“ durchsucht.⁹⁷²

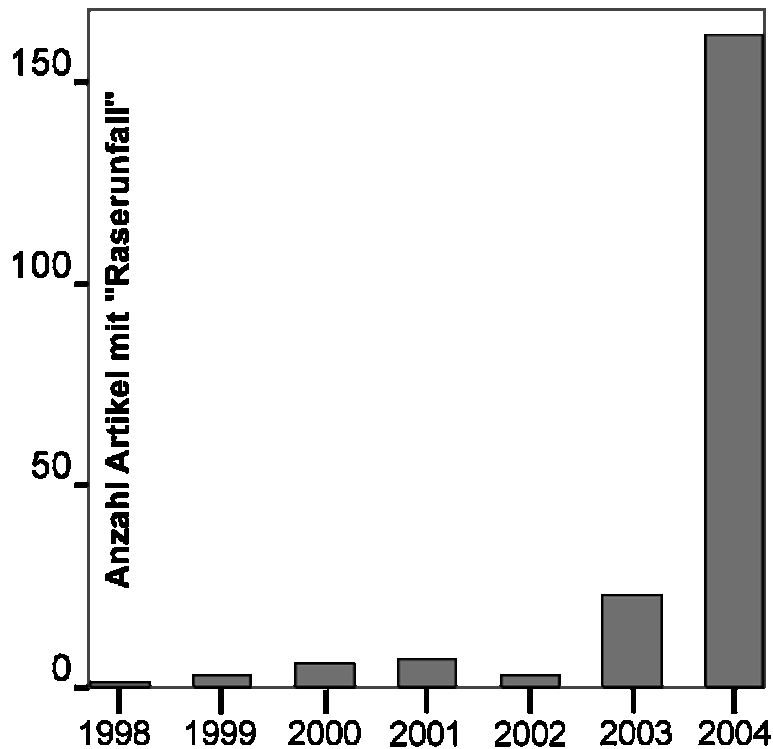


Abbildung 16: Berichterstattung über Raserunfälle

- 613 Die Resultate sind in der Abbildung 16 dargestellt. 1998 kam das Wort „Raserunfall“ in genau einem Bericht vor (im „Blick“ vom 11. Dezember 1998). 1999 benutzten drei Artikel das Wort „Raserunfall“, darunter eine Titelgeschichte des „Beobachters“, der auflagenstärksten Wochenzeitschrift der Schweiz,⁹⁷³ mit dem Titel: „Falsche Milde für Raser: Spektakuläre Raserunfälle fordern immer wieder Menschenleben. Doch die Gerichte lassen oft Milde walten. Zeit für eine andere Gangart“.⁹⁷⁴ Zwei Monate zuvor, am 3. September 1999, hatte der Unfall stattgefunden, den das Bundesgericht im BGE 130 IV 58 beurteilte. In den Jahren 2000 bis 2002 blieb die Anzahl der Nennungen von „Raserunfall“ bei sechs bis sieben pro Jahr. Im Jahr 2003 stieg sie auf 23, im Jahr 2004 auf 162 (!), wobei 144 dieser Nennungen nach dem 26. April 2004 (dem Tag, an dem BGE 130 IV 58 entschieden wurde) erfolgten.⁹⁷⁵

⁹⁷¹ Eine komplette Liste der erfassten Zeitungen und Zeitschriften ist unter www.swissdox.ch -> Porträt -> Medienliste erhältlich (besucht am 17. Februar 2005).

⁹⁷² Die Grenzen dieser Methode sind mir durchaus bewusst. Insbesondere kann es sein, dass vor 2003 über Raserunfälle berichtet wurde, ohne das Wort „Raserunfall“ zu verwenden.

⁹⁷³ Auflage gemäss Publicitas, MACH Basic 2003: 333'226, Leserzahl: 1'045'000. Eine höhere Auflage haben nur die Gratis-Kundenzeitungen der Lebensmittel-Grossverteiler.

⁹⁷⁴ Beobachter vom 29. Oktober 1999.

⁹⁷⁵ Eine kursorische Durchsicht der 162 Treffer für „Raserunfall“ 2004 ergab, dass praktisch keine Artikel in erster Linie über den Bundesgerichtsentscheid berichteten.

- 614 Auch diese exzessive Berichterstattung wird wieder zum Mittelwert regressieren. Dass durch den Medien-Hype der Eindruck entsteht, dass die auf Geschwindigkeitsexzesse zurückzuführenden Unfälle tatsächlich zugenommen haben, legt die Verfügbarkeitsheuristik (*availability heuristic*) nahe: gemäss der Verfügbarkeitsheuristik werden diejenigen Ereignisse als häufiger eingeschätzt, die einem leichter einfallen.⁹⁷⁶ Ereignisse, über die Massenmedien ständig berichten, sind leicht verfügbar, und ihre tatsächliche Häufigkeit wird überschätzt.⁹⁷⁷

4. Das „Autorennen“ Urteil des Bundesgerichts (BGE 130 IV 58)

- 615 Am 26. April 2004 hat das Bundesgericht in dem bereits auf S. 201 erwähnten Entscheid zwei Autofahrer, die sich auf öffentlichen Strassen ein Rennen geliefert und dabei einen tödlichen Unfall verursacht haben, auf eventualvorsätzliche Tötung erkannt und das vorinstanzliche Urteil, das die beiden Raser zu je 6,5 Jahren Gefängnis verurteilte, bestätigt (BGE 130 IV 58). Das Urteil wird von den Medien und anderen Kreisen, wie der Beratungsstelle für Unfallverhütung, ausdrücklich als Schritt in die richtige Richtung und Beitrag zur Erhöhung der Verkehrssicherheit begrüsst.⁹⁷⁸
- 616 Das Urteil lässt sich nur erklären, wenn man sich bewusst ist, dass die Höchststrafe für mehrfache fahrlässige Tötung drei Jahre Gefängnis beträgt.⁹⁷⁹ Diese Strafe erachtete das Bundesgericht, wie die beiden Vorinstanzen, offensichtlich als nicht angemessen. Dogmatisch lässt sich jedenfalls nicht begründen, warum ein Autofahrer einen Unfall eventualvorsätzlich verursachen soll; es handelt sich um einen Lehrbuch-Beispiel bewusster Fahrlässigkeit. Ein erstes (nie in der amtlichen Sammlung publiziertes) Raserurteil des Bundesgerichts aus dem Jahr 1986, das einen Sportwagenfahrer, der auf der Autobahn mit 240 km/h einen Unfall verursachte, wegen eventualvorsätzlicher Tötung verurteilte, wurde in der Lehre denn auch heftig kritisiert.⁹⁸⁰ Das Obergericht des Kantons Zürich folgt der neusten Rechtsprechung des Bundesgerichts nur mit ungutem Gefühl. Es weist darauf hin, dass es konsequenter wäre, den Strafraumen für fahrlässige Tötung zu erhöhen statt die Grenzen zwischen Vorsatz- und Fahrlässigkeitsdelikt zunehmend zu verwischen.⁹⁸¹
- 617 Warum das Bundesgericht auf eine hohe Strafe erkennen wollte, sagt das Urteil nicht. Es kann sein, dass das Bundesgericht annimmt, dass damit die Verkehrssicherheit gefördert

⁹⁷⁶ AMOS TVERSKY/DANIEL KAHNEMAN, Availability: A heuristic for judging frequency and probability, *Cognitive Psychology* 1973, 207-232.

⁹⁷⁷ Eine ausführlichere Analyse, wie sich Verfügbarkeit und soziale Mechanismen gegenseitig verstärken und zu suboptimalem Risikomanagement (auch Bestrafung von Rasern, wenn sie präventive Zwecke verfolgt, ist Risikomanagement) führen kann, präsentieren TIMUR KURAN/CASS R. SUNSTEIN, Controlling Availability Cascades, *Stanford Law Review* 1999, 683; Nachdruck in: SUNSTEIN (Hrsg.), 374-397.

⁹⁷⁸ Beratungsstelle für Unfallverhütung, FN 960.

⁹⁷⁹ Art. 117, 36, 68 Ziff. 1 StGB: weil zwei Menschen getötet wurden, ist die Strafe obligatorisch zu schärfen, aber der Richter ist gemäss Art. 68 Ziff. 1 StGB an das Höchstmass der Strafart (Gefängnis, daher drei Jahre, Art. 36 StGB) gebunden; er darf nach h.L. auf keine andere als die angedrohte Strafart erkennen, JÜRGEN-BEAT ACKERMANN, in: NIGGLI/WIPRÄCHTIGER, FN 959, Art. 68 N 35.

⁹⁸⁰ HANS SCHULTZ, Rechtsprechung und Praxis zum Strassenverkehrsrecht in den Jahren 1983-1987, *Bern* 1990, 94 f.; JEAN-PIERRE GUIGNARD, Note sur l'arrêt X., *JdT* 1988, 131.

⁹⁸¹ Tages-Anzeiger vom 17. Dezember 2004, 13.

wird, es kann auch sein, dass Sühneüberlegungen oder der grosse Druck der öffentlichen Meinung, die im Zeitpunkt des Urteils aufgrund mehrerer Unfälle so genannter „Balkanraser“ nach harten Strafen verlangte, eine Rolle spielte.

- 618 Ich kann mich des Eindrucks nicht erwehren, dass alle drei gerichtlichen Instanzen in diesem Fall der Illusion der abschreckenden Wirkung harter Strafen unterlegen sind. Dass es sich dabei um eine Illusion handelt, darf als erwiesen betrachtet werden. Ein Zusammenhang zwischen Strafhärte und Wiederholungsrate kann selbst bei Fahren in angetrunkenem Zustand nicht nachgewiesen werden. Es ist noch erheblich unwahrscheinlicher, dass die Bestrafung wegen eventualvorsätzlicher Tötung bei *Unfällen* eine abschreckende Wirkung hat. Fahren in angetrunkenem Zustand ist ein Vorsatzdelikt. Wer zuviel getrunken hat, weiss dies in der Regel. Wenn jemand trotzdem fährt, dann deshalb, weil er darauf vertraut, nicht erwischt zu werden. Raserunfälle sind – trotz gegenteiliger Auffassung des Bundesgerichts – Fahrlässigkeitsdelikte. Wer einen tiefergelegten VW Corrado VR 6 mit verdunkelter Heckscheibe und Doppelauspuff fährt, will sein Auto nicht bei einem Unfall verschrotten. Er vertraut – zweifellos fälschlicherweise, aber Selbstüberschätzung ist ein verbreitetes Phänomen (siehe hinten, S. 261 ff.) – darauf, dass er sein Fahrzeug beherrscht. Er wird sich daher nicht von der Strafe beeindrucken lassen, die für die Verursachung des Unfalls angedroht ist, weil er nicht mit dem Unfall rechnet. Er erkennt wohl, dass zu hohe Geschwindigkeit *generell* gefährlich ist, aber *ihm* wird natürlich nichts passieren, weil er sein Auto beherrscht. Wer Prävention betreiben will, muss die Wahrscheinlichkeit erhöhen, dass Geschwindigkeitsübertretungen bestraft werden. Die sich unter dem (falschen) Eindruck steigender Unfallzahlen und dem öffentlichen Druck drehende Spirale immer härterer Strafen bei SVG-Delikten, die offenbar auch vor dogmatischen Grenzen nicht Halt macht, muss gestoppt werden.

§ 14 Rückschaufehler (*hindsight bias*)

- 619 „Das musste ja so kommen“ – im Nachhinein erscheinen die Folgen einer Handlung oder der Ausgang einer Geschichte oft vorbestimmt. Die Tendenz, die Vorhersehbarkeit eines Ereignisses retrospektiv zu überschätzen, wird in der Psychologie als Rückschaufehler (*hindsight bias*) bezeichnet. Nach BARUCH FISCHHOFF.⁹⁸²

In der Rückschau überschätzen Menschen ständig, was vorhersehbar war. Nicht nur betrachten sie das, was geschehen ist, als unausweichlich, sondern sie meinen auch, dass es, bevor es geschehen ist, als „einigermassen unausweichlich“ erschien. Menschen meinen, dass andere den Ausgang eines Ereignisses viel besser hätten vorhersehen müssen, als dies tatsächlich der Fall war. Sie erinnern sich sogar falsch an ihre eigenen Vorhersagen, so dass sie im Nachhinein übertreiben, was sie vorher gewusst haben.

- 620 Der Rückschaufehler wurde erstmals 1975 durch FISCHHOFF systematisch untersucht.⁹⁸³ Seither wurde er in zahlreichen Studien erforscht und hat sich als eines der robustesten Phänomene der Gedächtnisforschung erwiesen. Eine umfassende Meta-Analyse von CHRISTENSEN-SZALANKSI/WILLHAM, die 128 Studien, publiziert in 40 Artikeln, umfasste, fand nur sechs Studien, in denen kein signifikanter Rückschaufehler festgestellt wurde.⁹⁸⁴ 2003 erschien eine Sondernummer der Zeitschrift *Memory*, die ausschliesslich dem Rückschaufehler gewidmet war⁹⁸⁵ und 2004 wurde eine weitere Meta-Analyse publiziert, die 95 Studien umfasste.⁹⁸⁶
- 621 Die Studien zum Rückschaufehler lassen sich nach ihrem Design in solche mit Gedächtnisdesign und solche mit hypothetischem Design einteilen.⁹⁸⁷ Beim Gedächtnisdesign beantworten die Versuchspersonen Fragen, ohne dass sie die richtige Antwort oder den Ausgang der Geschichte kennen, erfahren anschliessend (oder nach einer Verzögerung) die Antwort und werden dann (oder nach einer Verzögerung) gebeten, sich an ihre ursprüngliche Antwort zu erinnern. Beim hypothetischen Design werden die Versuchspersonen gebeten, einzuschätzen, wie vorhersehbar der Ausgang eines Szenarios für andere, oder für sich selbst, ohne das Wissen um den Ausgang gewesen wäre, oder wie viele Personen eine Frage richtig beantwortet hätten, wenn die Antwort nicht bekannt gewesen wäre.⁹⁸⁸ Der Rückschaufehler ist beim hypothetischen Design in der Regel grösser als beim Gedächtnisdesign.⁹⁸⁹

⁹⁸² BARUCH FISCHHOFF, For Those Condemned to Study the Past: Heuristics and Biases in Hindsight, in: KAHNEMAN/SLOVIC/TVERSKY (Hrsg.), 335-354, 341.

⁹⁸³ BARUCH FISCHHOFF, Hindsight ≠ Foresight: The Effect of Outcome Knowledge on Judgment Under Uncertainty, *Journal of Experimental Psychology* 1975, 288-299.

⁹⁸⁴ JAY J. CHRISTENSEN-SZALANKSI/CYNTHIA FOBIA WILLHAM, The Hindsight Bias: A Meta-Analysis, *Organizational Behavior and Human Decision Processes* 1991, 147-168.

⁹⁸⁵ *Memory* 2003, 329 ff.

⁹⁸⁶ REBECCA L. GUILBAULT/FRED B. BRYANT/JENNIFER HOWARD BROCKWAY/EMIL J. POSAVAC, A Meta-Analysis of Research on Hindsight Bias, *Basic and Applied Social Psychology* 2004, 103-117.

⁹⁸⁷ ULRICH HOFFRAGE/RÜDIGER F. POHL, Research on hindsight bias: A Rich Past, a Productive Present and a Challenging Future, *Memory* 2003, 329-335, 330.

⁹⁸⁸ RALPH HERTWIG/GERD GIGERENZER/ULRICH HOFFRAGE, The Reiteration Effect in Hindsight Bias, *Psychological Review* 1997, 194-202, 195 schlagen vor, den Begriff *hindsight bias* für das Phänomen zu

- 622 Nach dem verwendeten Stimulus-Material lassen sich die Studien einteilen in solche, die Allgemeinwissensfragen verwenden („Wie viele Knochen hat der menschliche Körper?“), solche, die Szenarien (Geschichten) verwenden (z.B. den Krieg zwischen den Briten und den Gurkas im 19. Jh. in Indien) und solche, die tatsächliche Ereignisse, über die in den Medien berichtet wird (z.B. Ergebnisse von Fussballspielen oder Bundestagswahlen) verwenden.
- 623 Während für die Gedächtnisforschung in erster Linie das Gedächtnisdesign interessant ist – die Studien in der Sondernummer von *Memory* folglich auch mehrheitlich dieses Design verwenden – ist für den Juristen die Kombination von hypothetischem Design mit Szenarien als Stimulus-Material von höchstem Interesse. Dieses Design reflektiert das, was ein Gericht macht, wenn es die Vorhersehbarkeit der Folgen einer Handlung im Nachhinein beurteilen muss.
- 624 Eng verwandt mit dem Rückschaufehler ist der Ausgangseffekt (*outcome effect*). Eine Entscheidung in einem Einzelfall sollte nach ihrem Potential, d.h. nach der Wahrscheinlichkeit, das sie zu einem positiven Resultat führt, beurteilt werden. Ob die Entscheidung richtig war, beurteilt sich unabhängig davon, ob das positive Resultat tatsächlich eingetreten ist. Ein Beispiel: Wenn zwei Personen die Wahl zwischen zwei Spielen haben, wobei (a) die Chance beinhaltet, Fr. 100 zu gewinnen, wenn eine faire Münze auf „Zahl“ fällt; und (b) die Chance, Fr. 100 zu gewinnen, wenn ein fairer Würfel eine sechs zeigt, dann ist die einzig rationale Entscheidung, das Spiel (a) zu wählen, da es einen höheren erwarteten Wert hat.⁹⁹⁰ Dies gilt auch dann, wenn die Person, die (b) gewählt hat, zufälligerweise gewinnt, während die Person, die (a) gewählt hat, verliert, weil die Münze auf Kopf fällt.
- 625 Bei einem Beispiel, bei dem sich die Wahrscheinlichkeiten einfach und mit Sicherheit berechnen lassen, leuchtet dies unmittelbar ein. Was aber, wenn die Situation ambivalent und unklar ist und verschiedene Indizien für eine grössere Wahrscheinlichkeit der einen oder anderen Folge der Entscheidung sprechen? Der Rückschaufehler führt in solchen Fällen dazu, dass die tatsächlichen Folgen der Entscheidung als wahrscheinlicher betrachtet werden, als sie es im Zeitpunkt der Entscheidung objektiv waren. Dementsprechend wird eine Entscheidung, die negative Konsequenzen hatte, eher als „falsch“ beurteilt, obwohl sie objektiv, im Zeitpunkt, in dem sie getroffen wurde, die richtige Entscheidung war. In einer Studie von JONATHAN BARON und JOHN C. HERSHEY wurde den Versuchspersonen gesagt, dass bei einer bestimmten Herzoperation acht von hundert Patienten an Komplikationen der Operation sterben. Die Versuchspersonen wurden gebeten, den Entscheid des Arztes zu beurteilen, die Operation durchzuführen. Die Versuchspersonen, denen gesagt wurde, dass die Operation erfolgreich verlaufen war, bewerteten die Entscheidung signifikant positiver als diejenigen, denen mitgeteilt wurde, dass der Patient an den Folgen der Operation

reservieren, dass beim Gedächtnisdesign beobachtet wird und den Ausdruck „*knew-it-all-along*“ Effekt für das beim hypothetischen Design beobachtete Phänomen zu verwenden. Da in der juristischen Literatur, und meinem Eindruck nach auch in den meisten psychologischen Publikationen, dieser terminologischen Unterscheidung nicht gefolgt wird, verzichte ich hier ebenfalls darauf.

⁹⁸⁹ Nachweise bei RÜDIGER F. POHL/MARKUS EISENHAEUER/OLIVER HARDT, SARA: A Cognitive Process Model to Simulate the Anchoring Effect and Hindsight Bias, *Memory* 2003, 337-356, 348.

⁹⁹⁰ RAANAN LIPSHITZ/DALYA BARAK, Hindsight wisdom: Outcome Knowledge and the Evaluation of Decisions, *Acta Psychologica* 1995, 105-125, 106; das Beispiel wird AMOS TVERSKY zugeschrieben.

starb.⁹⁹¹ Handlungen, die zu positiven Folgen führten, werden auch eher als im Einklang mit den im Zeitpunkt der Entscheidung geltenden Normen stehend gesehen als Handlungen mit negativen Folgen.⁹⁹² Grundsätzlich gilt: je schlimmer die Folgen einer Entscheidung, desto negativer wird das Verhalten bewertet, das zu den Folgen führte.⁹⁹³

- 626 Der Rückschaufehler wird allgemein als Nebenerscheinung adaptiven Lernens betrachtet. HAWKINS/HASTIE bezeichnen ihn als „die dunkle Seite erfolgreichen Lernens“.⁹⁹⁴ Selektive Aufmerksamkeit und selektive Abrufung von Erinnerungen sind in einer komplexen Umwelt notwendig. Um komplexe Entscheidungen zu treffen, entwickelt der Urteilende ein mentales Modell der Umwelt, das er aufgrund beobachteter Ereignisse ständig anpasst. Diese Anpassung des Modells an geänderte Verhältnisse ist eine Voraussetzung für induktives Lernen. Ist das Modell aber erst einmal an die neuen Umstände angepasst, ist es schwierig, noch zu rekonstruieren, wie der Entscheid gemäss dem „alten“ Modell erfolgt wäre.⁹⁹⁵ Solche Urteile aufgrund eines „veralteten“ Modells waren während eines grossen Teils der menschlichen Evolutionsgeschichte weder verlangt noch nützlich. Der Rückschaufehler steht daher durchaus im Einklang mit der evolutionären Sicht der Psychologie und ist weniger in die Kritik geraten als andere in der *heuristics and biases* Literatur beschriebenen Phänomene.
- 627 Im Folgenden wird ein Überblick über die psychologische Forschung zum Rückschaufehler gegeben, wobei nicht auf die verschiedenen – teilweise sehr komplexen – theoretischen Konzepte eingegangen wird.⁹⁹⁶ Anschliessend werden empirische Studien mit spezifisch rechtlichen Fragestellungen zusammengefasst und die eigene Studie dargestellt. Für die Literatur, die vor 1990 erschienen ist, verlasse ich mich auf den Übersichtsartikel von HAWKINS/HASTIE aus dem Jahr 1990.⁹⁹⁷

I. Psychologie: Empirische Studien

- 628 FISCHHOFF liess in seiner klassischen Studie die Wahrscheinlichkeit verschiedener möglicher Folgen eines historischen Ereignisses von einer Kontrollgruppe einschätzen, die keine Kenntnis von den tatsächlichen Folgen hatte. Andere Versuchspersonen, denen zuvor die angebliche Folge mitgeteilt worden war, sollten dieselben Einschätzungen treffen, und

⁹⁹¹ JONATHAN BARON/JOHN C. HERSHEY, Outcome Bias in Decision Evaluations, *Journal of Personality and Social Psychology* 1989, 569-579, 571 f. (1988).

⁹⁹² LISPHITZ/BARAK, FN 990, 115.

⁹⁹³ MICHAEL R. BALDWIN/CHRIS L. KLEINKE, Effects of Severity of Accident, Intent, and “Alcoholism is a Disease” Excuse on Judgments of a Drunk Driver, *Journal of Applied Social Psychology* 1994, 2097-2109. Weitere Nachweise bei PHILIP G. PETERS, Hindsight Bias and Tort Liability: Avoiding Premature Conclusions, *Arizona State Law Review* 1999, 1277-1314, 1283 Fn. 46 und RACHLINSKI, FN 190, 581 Fn. 36.

⁹⁹⁴ SCOTT A. HAWKINS/REID HASTIE, Hindsight: Biased Judgments of Past Events After the Outcomes Are Known, *Psychological Bulletin* 1990, 311-327, 323.

⁹⁹⁵ ULRICH HOFFRAGE/RALPH HERTWIG/GERD GIGERENZER, Hindsight Bias: A By-Product of Knowledge Updating?, *Journal of Experimental Psychology: Learning, Memory and Cognition* 2000, 566-581.

⁹⁹⁶ Einen Überblick zu diesen bieten die in der bereits erwähnten Sondernummer von *Memory* erschienen Artikel.

⁹⁹⁷ HAWKINS/HASTIE, FN 994.

zwar so, als ob sie die Folge nicht kennen würden. Es zeigte sich, dass die Versuchspersonen dazu nicht in der Lage waren und die Wahrscheinlichkeit der „tatsächlichen“ Folge systematisch überschätzten.⁹⁹⁸ Auch betrachteten die Versuchspersonen Informationen, die aufgrund der Folgen als relevant erschienen, als von vorneherein besonders relevant; dies im Unterschied zur Kontrollgruppe, die diesen Informationen keine vertiefte Bedeutung zumass.⁹⁹⁹ FISCHHOFF bezeichnete das Phänomen als „schleichenden Determinismus“ (*creeping determinism*), d.h. im Nachhinein erscheinen die Läufe der Welt vorbestimmt.

- 629 Explizite Warnungen, den Rückschaufehler nicht zu begehen, zeigen keinen oder nur einen sehr geringen Effekt. Auch wenn die Versuchspersonen darauf hingewiesen werden, dass Menschen den Rückschaufehler begehen und übertreiben, wie viel sie im Vorneherein gewusst hätten, begehen sie substantielle Rückschaufehler.¹⁰⁰⁰ GUILBAULT und Kollegen halten in ihrer Meta-Analyse fest, dass Manipulationen zur Reduktion des Rückschaufehlers i. d. R. keinen signifikanten Effekt haben.¹⁰⁰¹ Die einzige Manipulation, die einigermaßen Erfolg versprechend ist, ist, die Versuchspersonen zu zwingen, Gründe aufzuführen, die zu den Folgen geführt haben könnten, die *nicht* eingetreten sind.¹⁰⁰² Der Schuss kann aber nach hinten losgehen, wenn die Versuchspersonen gezwungen werden, sehr viele (zehn oder mehr) Gründe aufzuführen, warum es zu einem anderen Ergebnis hätte kommen können. Versuchspersonen empfinden dies als schwierig und nehmen daher vermehrt an, das „tatsächliche“ Ergebnis sei das einzige plausible gewesen.¹⁰⁰³
- 630 Eine detailliertere Schilderung des Szenarios (mehrere hundert Wörter) führt zu einem stärkeren Rückschaufehler.¹⁰⁰⁴
- 631 Der Rückschaufehler wurde auch mit visuellen Stimuli dokumentiert.¹⁰⁰⁵ Bei einem typischen Experiment werden den Versuchspersonen verschwommene Bilder gezeigt, die allmählich schärfer werden. Während die Kontrollgruppe nicht weiss, was auf den Bildern schliesslich zu sehen sein wird, weiss die Versuchsgruppe dies bereits im Voraus. Die Versuchsgruppe wird gefragt, wann jemand ohne Kenntnis vom Inhalt des Bildes den Inhalt frühestens erkennen kann. Dabei zeigt sich, dass dieser Zeitpunkt systematisch zu

⁹⁹⁸ FISCHHOFF, FN 983, 291.

⁹⁹⁹ FISCHHOFF, FN 983, 292.

¹⁰⁰⁰ BARUCH FISCHHOFF, Perceived Informativeness of Facts, *Journal of Experimental Psychology: Human Perception and Performance* 1978, 349-358, 354; RÜDIGER F. POHL/WOLFGANG HELL, No Reduction in Hindsight Bias after Complete Information and Repeated Testing, *Organizational Behaviour and Human Decision Processes* 1996, 49-58.

¹⁰⁰¹ GUILBAULT/ BRYANT/ BROCKWAY/POSAVAC, FN 986, 115.

¹⁰⁰² PAUL SLOVIC/BARUCH FISCHHOFF, On the Psychology of Experimental Surprises, *Journal of Experimental Psychology: Human Perception and Performance* 1977, 544-551.

¹⁰⁰³ LAWRENCE J. SANNA/NORBERT SCHWARZ/SHEVAUN L. STOCKER, When Debiasing Backfires: Accessible Content and Accessibility Experiences in Debiasing Hindsight, *Journal of Experimental Psychology* 2002, 497-502.

¹⁰⁰⁴ DONALD C. PENNINGTON, Being Wise After the Event: An Investigation of Hindsight Bias, *Current Psychological Research* 1981, 271-282.

¹⁰⁰⁵ ERIN M. HARLEY/KERI A. CARLSEN/GEOFFREY R. LOFTUS, The “Saw-It-All-Along” Effect: Demonstrations of Visual Hindsight Bias, *Journal of Experimental Psychology: Learning, Memory and Cognition* 2004, 960-968.

früh geschätzt wird.¹⁰⁰⁶ Explizite Warnungen, die den Rückschaufehler erklären und darauf hinweisen, dass er durch die Kenntnis des Inhaltes des Bildes verursacht wird, zeigen keine positive Wirkung.¹⁰⁰⁷ HARLEY/CARLSEN/LOFTUS weisen darauf hin, dass diese Ergebnisse Konsequenzen beispielsweise für Arzthaftungsprozesse haben können. Der Gutachter, der beurteilen muss, ob ein Tumor auf einem Röntgenbild vom behandelnden Arzt hätte erkannt werden können, weiss meist, dass ein Tumor vorhanden war und die Diagnose verpasst wurde. Es ist sehr schwierig bis unmöglich für ihn, diese Information bei der Beurteilung des Bildes zu ignorieren, was dazu führen kann, dass er die Erkennbarkeit des Tumors bejaht, obwohl er ohne das Wissen um seine Existenz nicht erkennbar gewesen ist.¹⁰⁰⁸

- 632 Persönlichkeitsmerkmale haben einen gewissen Einfluss auf den Rückschaufehler, können ihn aber nicht vollständig erklären. Personen, die ein starkes Bedürfnis nach positiver Selbstdarstellung haben, zeigen bei Studien mit Gedächtnisdesign einen verstärkten Rückschaufehler; d.h. sie geben häufiger an, dass sie die richtige Antwort auch gewusst hätten.¹⁰⁰⁹ Man darf annehmen, dass sie ungern zugeben, etwas nicht gewusst zu haben. Einen noch stärkeren Einfluss auf den Rückschaufehler zeigen Persönlichkeitsmerkmale, die unter den Begriffen „Intoleranz für Ambivalenz“ (*intolerance for ambiguity*) und Dogmatismus zusammengefasst werden. Menschen, die diese Persönlichkeitsmerkmale aufweisen, haben ein verstärktes Bedürfnis nach Sicherheit, Klarheit und einer geordneten, sinnhaften und vorhersehbaren Welt.¹⁰¹⁰ Sie zeigen ebenfalls einen verstärkten Rückschaufehler.¹⁰¹¹ Von den „grossen fünf“ Persönlichkeitsmerkmalen – Extraversion, Neurotizismus, Freundlichkeit, Gewissenhaftigkeit, Offenheit für Erfahrung – scheint am ehesten Gewissenhaftigkeit mit einem verstärkten Rückschaufehler assoziiert zu sein.¹⁰¹² Persönlichkeitsmerkmale haben einen stärkeren Einfluss in Studien mit hypothetischem Design als in Studien mit Gedächtnisdesign.¹⁰¹³
- 633 Eine Umkehr des Rückschaufehlers kann bei sehr überraschenden Resultaten auftreten.¹⁰¹⁴ In diesem Fall tritt ein Phänomen auf, dass man mit „Also damit konnte nun wirklich keiner rechnen“ zusammenfassen lässt und dazu führt, dass der „tatsächliche“ Ausgang im Nachhinein als sehr unwahrscheinlich bewertet wird. Wichtig ist, dass die Umkehrung nur bei überraschendem Ausgang, der (fast) keinen Sinn ergibt, nicht aber bei wenig oder einigermaßen überraschendem, erklärbarem Ausgang auftritt.¹⁰¹⁵ Weder ein Rückschau-

¹⁰⁰⁶ HARLEY/CARLSEN/LOFTUS, FN 1005, 965.

¹⁰⁰⁷ HARLEY/CARLSEN/LOFTUS, FN 1005, 963.

¹⁰⁰⁸ HARLEY/CARLSEN/LOFTUS, FN 1005, 966.

¹⁰⁰⁹ JOCHEN MUSCH, *Personality Differences in Hindsight Bias*, Memory 2003, 473-489, 485.

¹⁰¹⁰ MUSCH, FN 1009, 477.

¹⁰¹¹ MUSCH, FN 1009, 484.

¹⁰¹² MUSCH, FN 1009, 485.

¹⁰¹³ MUSCH, FN 1009, 486.

¹⁰¹⁴ CHEZY OFIR/DAVID MAZURSKY, *Does a Surprising Outcome Reinforce or Reverse the Hindsight Bias?*, Organizational Behavior and Human Decision Making Processes 1997, 51-57.

¹⁰¹⁵ MARK V. PEZZO, *Surprise, Defence Or Making Sense: What Removes Hindsight Bias?*, Memory 2003, 421-441.

fehler noch seine Inversion treten auf, wenn ein Ereignis gänzlich den Erwartungen entspricht und überhaupt nicht überraschend ist.¹⁰¹⁶

- 634 Die Umkehr tritt vor allem bei überraschendem Ausgang auf, der ein schlechtes Licht auf die Entscheidungen und Fähigkeiten des Handelnden wirft und somit sein Selbstwertgefühl beeinträchtigt. In einem Experiment von MELVIN M. MARK und Kollegen nahmen die Versuchspersonen an einem Börsenspiel teil. Nach der dritten Woche wurde einem Teil der Versuchspersonen mitgeteilt, dass der Kurs der von ihnen gekauften Aktien leider sehr stark gesunken sei. Anschliessend wurden die Versuchspersonen, die die Aktien gekauft hatten, ihre Gegenspieler und neutrale Beobachter gefragt, ob der Kursverlust vorhersehbar gewesen wäre. Die Käufer der Aktien beurteilten den Kursverlust als weniger vorhersehbar als Gegenspieler und Beobachter, für die der Kursverlust keine Auswirkungen auf das Selbstwertgefühl hatte. Für keinen anderen Ausgang wurde die Vorhersehbarkeit von Spieler, Gegenspieler und Beobachter unterschiedlich beurteilt.¹⁰¹⁷ Dieses Ergebnis reflektiert eine oft angetroffene Situation in Haftpflichtprozessen: der Beklagte taxiert die negativen Folgen seiner Handlung als absolut nicht voraussehbar, während Kläger und Gericht anderer Auffassung sind.

II. Psychologie: Empirische Studien mit Experten

A. Ärzte

- 635 ARKES/WORTMANN/SAVILLE/HARKNESS verglichen die Diagnose durch Ärzte, die eine nicht-diagnostizierte Fallbeschreibung lasen, mit der Diagnose durch Ärzte, die dieselbe Fallbeschreibung beurteilten, die aber bereits mit einer bestimmten Diagnose versehen war. Beide Gruppen wurden gebeten anzugeben, für wie wahrscheinlich sie eine von vier zur Auswahl stehenden Diagnosen hielten. Die Versuchsgruppe wurde gebeten, diese Einschätzung ungeachtet der bereits erfolgten Diagnose zu machen. Das Experiment zeigte den üblichen Rückschaufehler, d.h. die Ärzte in der Versuchsgruppe überschätzten die Wahrscheinlichkeit, mit der sie die tatsächliche Diagnose ohne vorgängige Kenntnis der Diagnose getroffen hätten. Der Rückschaufehler ist also auch dann robust, wenn die Urteile über ein Szenario abgegeben werden, über das die befragten Personen sehr viel wissen.¹⁰¹⁸ ARKES und Kollegen weisen darauf hin, dass es ihr Ergebnis als angezeigt erscheinen lässt, bei der Einholung einer Zweitmeinung dem zweiten Arzt die Diagnose des ersten Arztes nicht bekannt zu geben, wenn immer dies möglich ist.
- 636 In einer weiteren Studie testeten ARKES und Kollegen die von SLOVIC/FISCHHOFF vorgeschlagene Strategie zur Reduktion des Rückschaufehlers. Erneut wurde den Ärzten eine Fallbeschreibung vorgelegt und sie anschliessend gebeten, eine Diagnose abzugeben. Bei der Versuchsgruppe war dem Fall bereits eine der drei Diagnosen Alkoholentzug,

¹⁰¹⁶ Pezzo, FN 1015, 427.

¹⁰¹⁷ MELVIN M. MARK/RENEE REITER BOBURKA/KRISTEN M. EYSELL/LAURIE L. COHEN/STEVEN MELLOR, „I Couldn't Have Seen It Coming“: The Impact of Negative Self-Relevant Outcomes on Retrospections About Foreseeability, *Memory* 2003, 443-454.

¹⁰¹⁸ HAL R. ARKES/ROBERT L. WORTMANN/PAUL D. SAVILLE/ALLAN R. HARKNESS, Hindsight Bias Among Physicians Weighing the Likelihood of Diagnoses, *Journal of Applied Psychology* 1981, 252-254, 254.

Alzheimer oder Hirnverletzung zugeordnet. Der Vergleich von Kontroll- und Testgruppe zeigte den typischen Rückschaufehler, d.h. die Ärzte in der Versuchsgruppe gaben auf die Frage, welche Wahrscheinlichkeit sie jeder Diagnose in Unkenntnis der tatsächlichen Diagnose zugeordnet hätten, Antworten, die systematisch in Richtung der „tatsächlichen“ Diagnose verzerrt waren. Wurden die Versuchspersonen aber gezwungen, je einen aus der Fallbeschreibung ersichtlichen Grund für jede der drei möglichen Diagnosen aufzuführen, verschwand der Rückschaufehler weitgehend.¹⁰¹⁹

B. Buchprüfer

- 637 CRAIG EMBY, ALEXANDER GELARDI und JORDAN LOWE baten 122 Partner der damals noch *Big Five* Buchprüfungsfirmen die finanzielle Lage eines fiktiven Unternehmens und deren Beurteilung durch einen anderen Buchprüfer zu bewerten. Wurde den Partnern gesagt, dass die Gesellschaft inzwischen Konkurs gegangen sei, betrachteten sie diesen Ausgang als vorhersehbarer als diejenigen Partner, denen keine Informationen über das Schicksal der Firma mitgeteilt worden waren. Auch bewerteten sie Indizien, die für einen baldigen Konkurs sprachen, als wichtiger als Indizien, die dagegen sprachen (es gab gleich viele Indizien, die für oder gegen einen Konkurs sprachen).¹⁰²⁰
- 638 EMBY und Kollegen fanden auch einen signifikanten Ausgangseffekt, d.h. die Beurteilung der Revision hing vom positiven oder negativen Ausgang des Szenarios ab (die Qualität der Revision wurde auf einer Skala von 1-10 bei gutem Ausgang mit 6,9 bewertet, bei schlechtem Ausgang mit 5,6 und ohne Angabe eines Ausgangs mit 6,6).¹⁰²¹ Die Resultate von EMBLY und Kollegen decken sich weitgehend mit den Resultaten einer Studie von ANDERSON, LOWE und RECKERS aus dem Jahr 1993.¹⁰²² ANDERSON und Kollegen stellten insbesondere auch fest, dass ihre Versuchspersonen in Abhängigkeit vom Ausgang Indizien, die auf einen positiven oder negativen Ausgang hindeuteten, unterschiedliches Gewicht beimessen.¹⁰²³

C. Richter

- 639 GUTHRIE /RACHLINSKI/WISTRICH legten amerikanischen *magistrate court judges* eine kurze Zusammenfassung eines erstinstanzlichen Urteils vor. Einem Drittel der Richter wurde gesagt, dass das Berufungsgericht das Urteil aufgehoben und mit der Auflage, eine geringere Sanktion auszusprechen, an die Vorinstanz zurückgewiesen habe, einem Drittel, dass das Urteil in der Berufung bestätigt wurde und einem Drittel, dass das Urteil und die vom erstinstanzlichen Gericht ausgesprochenen Sanktionen kassiert wurde. Die Richter

¹⁰¹⁹ ARKES/FAUST/GUILMETTE/HART, Eliminating the Hindsight Bias, *Journal of Applied Psychology* 1988, 305-307, zitiert nach HAWKINS/HASTIE, FN 997, 318.

¹⁰²⁰ CRAIG EMBLY/ALEXANDER M. G. GELARDI/D. JORDAN LOWE, A Research Note on the Influence of Outcome Knowledge on Audit Partners' Judgments, *Behavioral Research in Accounting* 2002, 86-103.

¹⁰²¹ EMBLY/GELARDI/LOWE, FN 1020, 97.

¹⁰²² JOHN C. ANDERSON/D. JORDAN LOWE/PHILIP M.J. RECKERS, Evaluation of auditor decisions: Hindsight bias effects and the expectation gap, *Journal of Economic Psychology* 1993, 711-737.

¹⁰²³ ANDERSON/LOWE/RECKERS, FN 1022, 730.

Besonderer Teil

wurden gefragt, welchen Ausgang des Berufungsverfahrens sie selber, in Unkenntnis des tatsächlichen Ausgangs, für wie wahrscheinlich gehalten hätten.¹⁰²⁴

- 640 GUTHRIE und Kollegen stellten einen signifikanten Rückschaufehler fest. Beispielsweise gaben 81,5 % der Richter, denen gesagt worden war, dass das Urteil vom Berufungsgericht bestätigt worden war, an, dass sie diesen Ausgang des Berufungsverfahrens vorausgesehen hätten, gegenüber nur 27,8 % der Richter, denen gesagt wurde, dass das Urteil aufgehoben wurde.¹⁰²⁵
- 641 Auch die 59 bundes- und gliedstaatlichen amerikanischen Richter, die an einer Studie von ANDERSON und Kollegen teilnahmen, begingen substantielle Rückschaufehler. Sie bewerteten die Qualität der Arbeit der Buchprüfer zudem je nach positivem oder negativem Ausgang unterschiedlich und unterlagen somit einem Ausgangseffekt.¹⁰²⁶
- 642 REID HASTIE und W. KIP VISCUSI berichten von einer eigenen Studie, in der Richter dem Rückschaufehler weniger stark unterlagen als Laien und schlugen vor, die Zusprechung von Strafschadenersatz der Zuständigkeit der Geschworenen zu entziehen und in die Hände der Richter zu legen.¹⁰²⁷ Wie LEMPERT jedoch überzeugend nachweist, erlauben die von HASTIE/VISCUSI erhobenen Daten weder den von ihnen gezogenen Schluss auf die geringere Fehleranfälligkeit von Richtern noch stützen sie die von ihnen empfohlene Reform des Prozessrechts.¹⁰²⁸

III. Recht: Empirische Studien

- 643 KIM A. KAMIN und JEFFREY RACHLINSKI untersuchten, ob die Tatsache, dass eine Flut mit Sachschäden stattgefunden hat, die Einschätzung beeinflusst, ob die politische Gemeinde, die den Schaden erlitten hat, Massnahmen hätte treffen müssen, um Schaden zu vermeiden.¹⁰²⁹ In der Vorschau-Perspektive nahmen die Versuchspersonen – Studierende der Universität Stanford – an einer fiktiven Sitzung des Gemeinderates teil. Sie hörten (ab Band) die Aussagen von sechs Sachverständigen zur Wahrscheinlichkeit und den Folgen einer Flut auf Gemeindegebiet sowie den Kosten von Schutzmassnahmen. Anschliessend wurden sie gefragt, ob sie der Gemeinde empfehlen würden, Massnahmen gegen die Folgen einer Flut zu treffen.

¹⁰²⁴ GUTHRIE/RACHLINSKI/WISTRICH, FN 20, 802.

¹⁰²⁵ GUTHRIE/RACHLINSKI/WISTRICH, FN 20, 802.

¹⁰²⁶ ANDERSON/LOWE/RECKERS, FN 1022, 730. Siehe auch JOHN C. ANDERSON/MARIANNE JENNINGS/D. JORDAN LOWE/PHILIP M.J. RECKERS, The Mitigation of Hindsight Bias in Judges' Evaluations of Auditor Decisions, *Auditing: Journal of Practice and Theory* 1997, 20-39.

¹⁰²⁷ REID HASTIE/W. KIP VISCUSI, What Juries Can't Do Well: The Jury's Performance as a Risk Manager, *Arizona Law Review* 1998, 901-921.

¹⁰²⁸ RICHARD LEMPERT, Juries, Hindsight, And Punitive Damage Awards: Failures of a Social Science Case for Change, *DePaul Law Review* 1999, 867-893. Siehe auch JEFFREY RACHLINSKI, Heuristics and Biases in the Courts: Ignorance or Adaptation?, *Oregon Law Review* 2000, 61-102, 65, der darauf hinweist, dass kognitive Täuschungen einfacher bei anderen zu entdecken sind als bei einem selbst und daher eher bei Geschworenen als bei Richtern erkannt und korrigiert werden.

¹⁰²⁹ KIM A. KAMIN/JEFFREY RACHLINSKI, Ex Post ≠ Ex Ante: Determining Liability in Hindsight, *Law and Human Behavior* 1995, 89-104.

- 644 Die Versuchspersonen in der Rückschau-Gruppe hörten dieselben sechs Zeugen an; ihnen wurde aber gesagt, dass sie als Geschworene in einem Prozess über die Haftung der Stadt zu entscheiden hätten, nachdem eine Flut das Gebäude des Klägers zerstört hatte.¹⁰³⁰
- 645 Nur 24 % der Versuchspersonen in der Vorschau-Gruppe empfahlen der Gemeinde, Massnahmen gegen die Folgen einer Flut zu treffen, aber 57 % der Personen in der Rückschau-Gruppe beurteilten den Entscheid, keine Massnahmen zu treffen, als fahrlässig und bejahten die Haftung der Gemeinde für die Flutschäden.¹⁰³¹ Die bei einem Teil der Rückschau-Gruppe eingefügte Instruktion, sich vorzustellen, wie die Ereignisse, die zu den Flutschäden geführt haben, sich anders abgespielt haben könnten, hatte keinen Einfluss auf die Beurteilung der Haftung.¹⁰³² Erfolgreiches *De-biasing* scheint vorauszusetzen, dass die Versuchspersonen tatsächlich aktiv Gründe für und gegen den Eintritt des Ereignisses aufführen; die bloße Aufforderung dazu genügt nicht.¹⁰³³
- 646 SUSAN und GARY LABINE legten ihren Versuchspersonen – aus der Gesamtbevölkerung rekrutiert – die Anamnese eines psychisch kranken Patienten und eine Beschreibung der Behandlungsschritte des Therapeuten vor. Einem Drittel der Versuchspersonen wurde gesagt, dass der Patient in der Folge gewalttätig wurde, einem Drittel, dass der Patient nicht gewalttätig wurde und einem Drittel wurde kein Ausgang mitgeteilt.¹⁰³⁴
- 647 Nach dem Tarasoff-Urteil (*Tarasoff v. Regents of University of California*, 1976) muss ein Therapeut, der die Gefährlichkeit eines Patienten erkennen kann, vernünftige Sorgfalt („*reasonable care*“) anwenden, um zu verhindern, dass der Patient Dritte schädigt. Die Versuchspersonen, denen mitgeteilt wurde, dass der Patient gewalttätig wurde, beurteilten die Behandlung weniger positiv als die Kontrollgruppe und schätzten die Wahrscheinlichkeit der Gewalttätigkeit höher ein (78 % gegenüber 51 % in der „keine Gewalt“ und 62 % in der „ohne Ausgang“ Gruppe). 24 % der Versuchspersonen, denen mitgeteilt wurde, dass der Patient gewalttätig wurde, hielten den Therapeuten für fahrlässig, gegenüber nur 9 % in der „ohne Ausgang“ und 6 % in der „keine Gewalt“ Gruppe.¹⁰³⁵
- 648 In einer Studie von HASTIE/SCHADKE/PAYNE wurde getestet, ob die vorausschauende oder zurückblickende Perspektive einen Einfluss auf die Zusprechung von Strafschadenersatz (*punitive damages*) hat.¹⁰³⁶ Nur 33 % der Versuchspersonen waren vorausschauend der Meinung, dass eine Eisenbahngesellschaft ohne Verbesserung der Sicherheit eines Streckenabschnitts die Strecke nicht mehr länger befahren dürfe, aber 67 % der Versuchspersonen

¹⁰³⁰ Nach der „*Learned Hand*“ Formel ist die Haftung dann gegeben, wenn Kosten der Schutzmassnahmen \leq Schadenssumme \times Eintretenswahrscheinlichkeit (*United States v. Carroll Towing Co.* (1947)). Die Haftung hängt also entscheidend davon ab, wie die Versuchspersonen die Wahrscheinlichkeit des Schadenseintritts beurteilen.

¹⁰³¹ KAMIN/RACHLINSKI, FN 1029, 98.

¹⁰³² KAMIN/RACHLINSKI, FN 1029, 98.

¹⁰³³ KAMIN/RACHLINSKI, FN 1029, 100.

¹⁰³⁴ SUSAN J. LABINE/GARY LABINE, Determination of Negligence and the Hindsight Bias, *Law and Human Behavior* 1996, 501-516.

¹⁰³⁵ LABINE/LABINE, FN 1034, 510.

¹⁰³⁶ REID HASTIE/DAVID A. SCHADKE/JOHN W. PAYNE, Juror Judgments in Civil Cases: Hindsight Effects on Judgments of Liability for Punitive Damages, *Law and Human Behavior* 1999, 597-613.

sonen hielten das Befahren der Strecke ohne zusätzliche Sicherheitsmassnahmen für grobfahrlässig (*reckless*), nachdem es zu einem Unfall gekommen war, und sprachen Strafschadenersatz zu.¹⁰³⁷

- 649 Nach amerikanischem Zivilrecht hat eine Person, deren Wohnung zu Unrecht von der Polizei durchsucht wurde, unabhängig vom Erfolg der Suche einen Anspruch auf Schadenersatz. Ob eine Durchsuchung zu Recht erfolgt, hängt unter anderem davon ab, ob die Polizei einen Durchsuchungsbefehl vorweisen kann oder ob der Durchsuchung zugestimmt wurde. Der Durchsuchung zustimmen kann einerseits der Durchsuchte, aber unter gewissen Umständen auch eine Drittperson, wenn der Dritte – vereinfacht ausgedrückt – Mitbesitzer der Wohnung ist.¹⁰³⁸ Wann ein Dritter berechtigt ist, der Durchsuchung zuzustimmen, ist nicht einfach zu entscheiden.
- 650 KAGEHIRO und Kollegen legten ihren Versuchspersonen die Beschreibung eines Falles vor, bei dem die Polizei die Wohnung eines Verdächtigen, die dieser zusammen mit einem Dritten bewohnte, ohne Zustimmung des Verdächtigen durchsuchte. Gewisse Zimmer wurden gemeinsam genutzt, andere Zimmer hingegen nicht, was es nach dem anwendbaren Fallrecht sehr schwierig macht zu entscheiden, ob der Dritte der Durchsuchung der gesamten Wohnung zustimmen darf. KAGEHIRO und Kollegen variierten verschiedene Aspekte des Falles, den sie ihren Versuchspersonen vorlegten.¹⁰³⁹ Für die vorliegende Darstellung relevant ist, dass 92 % der Versuchsperson das Recht der Drittperson bejahten, der Durchsuchung der gemeinsam bewohnten Räume zuzustimmen, wenn der Verdächtige zufällig abwesend war und bei der Durchsuchung belastendes Beweismaterial gefunden wurde, aber nur 42 %, wenn kein belastendes Material gefunden wurde.¹⁰⁴⁰ Die Tatsache, dass belastende Indizien gefunden wurden, hat einen Einfluss auf eine völlig andere Frage, die rechtlich gesehen ungeachtet des Ausgangs der Suche gleich zu beurteilen ist.¹⁰⁴¹
- 651 Gemäss einer Studie von KATHRYN KALDOUS hängt die Beurteilung des Verschuldens eines Buchprüfers, der übersieht, dass gewisse Aktiven der Bilanz überbewertet sind, in erster Linie von der Schwere der Folgen ab (Konkurs und Verlust zahlreicher Stellen in einem Fall; Überschuldung, aber Sanierung im anderen Fall). Die Qualität der Buchprüfungsarbeit hat, wenn die Folgen des Versehens schwer sind, keinen Einfluss auf die Haftungsfrage; der sorgfältige Buchprüfer haftet mit anderen Worten in diesen Fällen genau so häufig wie der unsorgfältige.¹⁰⁴²

¹⁰³⁷ HASTIE/SCHADKE/PAYNE, FN 1036, 605.

¹⁰³⁸ DOROTHY K. KAGEHIRO/RALPH B. TAYLOR/WILLIAM S. LAUFER/ALAN T. HARLANDT, Hindsight Bias and Third-Party Consentors to Warrantless Police Searches, *Law and Human Behavior* 1991, 305-314.

¹⁰³⁹ So z.B. ob der Durchsuchte anwesend war und gegen die Durchsuchung protestierte als die Drittperson der Durchsuchung zustimmte oder zufällig abwesend war.

¹⁰⁴⁰ KAGEHIRO/TAYLOR/LAUFER/HARLANDT, FN 1038, 310.

¹⁰⁴¹ Für eine weitere Studie zu Rückschaufehlern bei der Beurteilung von Hausdurchsuchungen siehe JONATHAN D. CASPER/KENNETTE BENEDICT/JO L. PERRY, Juror Decision Making, Attitudes, and the Hindsight Bias, *Law and Human Behavior* 1989, 291-310.

¹⁰⁴² KATHRYN KADOUS, The Effects of Audit Quality and Consequence Severity on Juror Evaluations of Auditor Responsibility for Plaintiff Losses, *The Accounting Review* 2000, 327-341.

IV. Eigene Studie zum Rückschaufehler

A. Frage zum Rückschaufehler (2003)

652 Den Richtern der Zivil- und Strafgerichte der Kantone Aargau, St. Gallen und Zürich wurde folgender Sachverhalt zur Beurteilung vorgelegt (2003):

Vergegenwärtigen Sie sich noch einmal den Sachverhalt der vorangehenden Frage (Unfall wegen Missachtung des Rechtsvortritts).¹⁰⁴³

Wie sich im Strafverfahren herausstellt, wurde der fehlbare Lastwagenfahrer bereits vor sieben Jahren wegen Fahrens in angetrunkenem Zustand verurteilt. Zu einem Unfall kam es damals nicht. Der Lastwagenfahrer wurde damals mittels Strafbefehl zu einer bedingten Gefängnisstrafe von 14 Tagen verurteilt.

Soweit Sie dies beurteilen können, bereut der Lastwagenfahrer seine Tat aufrichtig. Eine Therapie wegen Alkoholsucht lehnt er ab, da es sich, wie er sagt, um zwei isolierte Einzelfälle handle und er kein Alkoholiker sei.

Das erstinstanzliche Gericht verurteilt den Lastwagenfahrer wegen fahrlässiger schwerer Körperverletzung, Fahrens in angetrunkenem Zustand und grober Verkehrsregelverletzung zu einer bedingten Gefängnisstrafe von vier Monaten.

653 Der „tatsächliche“ Ausgang des Berufungsverfahrens wurde wie folgt variiert:

- Gruppe „Strafreduktion, unbedingt“: Das Berufungsgericht bestätigt das Urteil des erstinstanzlichen Gerichts im Schuldpunkt. Im Strafpunkt reduziert es die Dauer der Gefängnisstrafe auf drei Monate, ordnet aber den unbedingten Vollzug an.
- Gruppe „Strafreduktion, bedingt“: Das Berufungsgericht bestätigt das Urteil des erstinstanzlichen Gerichts im Schuldpunkt. Im Strafpunkt senkt es die Dauer der Gefängnisstrafe auf drei Monate und gewährt den bedingten Vollzug.
- Gruppe „Bestätigung“: Das Berufungsgericht bestätigt das Urteil des erstinstanzlichen Gerichts in Schuld- und Strafpunkt vollumfänglich.

654 Anschliessend wurden die Richterinnen und Richter gefragt, welcher Ausgang des Berufungsverfahrens aufgrund des geschilderten Sachverhaltes am wahrscheinlichsten gewesen sei. Es standen die drei geschilderten Möglichkeiten (Strafreduktion und unbedingter Vollzug; Strafreduktion und bedingter Vollzug; vollumfängliche Bestätigung) zur Auswahl.

655 Die Hypothese ist, dass sich die Richter durch das „tatsächliche“ Urteil des Berufungsgerichts beeinflussen lassen und denjenigen Ausgang des Berufungsverfahrens als wahrscheinlicher einschätzen, der dem „tatsächlichen“ Ausgang entspricht.

B. Resultate (2003)

656 Die Resultate sind in der Tabelle 16 dargestellt. Die Antworten, die aufgrund der Hypothese häufiger zu erwarten sind, sind fett gedruckt (auf Hundert fehlende Prozent sind auf Rundungsdifferenzen zurückzuführen).

¹⁰⁴³ Es handelt sich dabei um den beim Ankereffekt, S. 82, verwendeten Sachverhalt.

Gruppe ↓	Antworten			Total
	Strafreduktion; unbedingt	Strafreduktion; bedingt	Bestätigung	
Strafreduktion; unbedingt	24 (44%)	5 (9%)	25 (46%)	54
Strafreduktion; bedingt	15 (24%)	21 (33%)	27 (43%)	63
Bestätigung	22 (36%)	10 (16%)	28 (47%)	60
Total	61 (34%)	36 (20%)	80 (45%)	177

Tabelle 16: Rückschaufehler: Resultate der Umfrage 2003

- 657 177 Richterinnen und Richter haben die Frage beantwortet; 54 der Gruppe „Strafreduktion, unbedingt“, 63 der Gruppe „Strafreduktion, bedingt“ und 60 der Gruppe „Bestätigung“. Alles in allem wurde eine Reduktion und des Strafmasses bei gleichzeitiger Gewährung des bedingten Vollzugs als am wenigsten wahrscheinlich und die vollumfängliche Bestätigung des erstinstanzlichen Urteils als am wahrscheinlichsten betrachtet.
- 658 Bei den Antworten zeigte sich ein wesentlicher Einfluss des „tatsächlichen“ Ausgangs des Berufungsverfahrens. Ein Drittel der Richter, denen mitgeteilt wurde, dass das Berufungsgericht den bedingten Strafvollzug gewährt hatte, betrachteten diesen Ausgang des Berufungsverfahrens als am wahrscheinlichsten, gegenüber nur 9 % resp. 16 %, denen gesagt wurde, dass das Berufungsgericht die Strafe reduziert und den unbedingten Vollzug angeordnet resp. das Urteil der Erstinstanz vollumfänglich bestätigt habe. Bei allen Gruppen wurde der „tatsächliche“ Ausgang des Verfahrens als wahrscheinlicher beurteilt als in den anderen Gruppen. Die Resultate sind statistisch signifikant auf dem 5 % Level.¹⁰⁴⁴

C. Frage zum Rückschaufehler (2004)

- 659 Den Richtern der Zivil- und Strafgerichte der Kantone beider Basel, Bern und Graubünden wurde folgender Sachverhalt zur Beurteilung vorgelegt (2004):

Bei der Erstellung von Mietshäusern liess der verantwortliche Bauleiter B. Aushubmaterial in einem Tälchen ablagern. Das Aushubmaterial staute den durch das Tälchen fliessenden Bach. Der flache Teich, den der Bach vorher gebildet hatte, wurde dadurch viel grösser und tiefer (die Tiefe vergrösserte sich von 25 cm auf 4 m).

- 660 Bei der Hälfte der Fragebogen wurden folgende Angaben zu den Folgen der Handlungen des Bauleiters gemacht:

Im Februar, als der Teich teilweise zugefroren war, sahen die 7 bzw. 11 Jahre alten Brüder Luca und Marco einen verlassenen Ball auf dem Eis liegen. Der jüngere Knabe betrat das Eis, um den Ball zu holen. Dabei brach er im Eis ein und ertrank zusammen mit seinem Bruder, der ihm zu Hilfe kommen wollte.

- 661 Es handelt sich dabei um den leicht gekürzten Sachverhalt aus BGE 83 II 89 = Pra. 57 Nr. 21. Da der Bundesgerichtsentscheid im Original auf Italienisch veröffentlicht wurde, nicht als Leitentscheid gilt und aus dem Jahr 1957 stammt, ist anzunehmen, dass die meisten der befragten (deutschsprachigen) Richterinnen und Richter den Fall nicht kannten.

¹⁰⁴⁴ $\chi^2 = 12,727$; $p = 0,013$.

- 662 Die Hälfte der Fragebogen jeder Gruppe waren mit einem Bild des Bauleiters in Krawatte und Anzug versehen; auf der anderen Hälfte der Fragebogen trägt der Bauleiter ein weisses, kragenloses Sweatshirt. Das Bild diente zwei Zwecken; einerseits sollte damit die Hypothese getestet werden, dass ein ordentlicher, Anzug tragender Bauleiter weniger häufig als fahrlässig angesehen wird als ein leger gekleideter; zweitens diente das Bild zur Ablenkung von der Manipulation in der Fragestellung zum Hofeffekt. Es wäre einigermaßen auffällig gewesen, wenn nur bei einer Frage ein Bild verwendet worden wäre. Der Verdacht, dass das Bild der Manipulation dient, liegt dann nahe.
- 663 Die immer gleich lautende Frage war, ob der Bauleiter fahrlässig gehandelt habe, weil er den Teich weder durch eine Warntafel noch durch eine Abschränkung sicherte.
- 664 Die Hypothese ist, dass die Kenntnis der Folgen zu einer häufigeren Bejahung der Pflichtwidrigkeit führen würde.

D. Resultate (2004)

- 665 Eine grosse Mehrheit der befragten Richterinnen und Richter bejahten wie das Bundesgericht die Fahrlässigkeit des Bauleiters. Die – schwache – Manipulation, ob der Bauleiter auf dem Foto einen Anzug und Krawatte trug, zeigte keine Wirkung.¹⁰⁴⁵

Gruppe ↓		Antworten		
		Fahrlässigkeit bejaht	Fahrlässigkeit verneint	Total
Mit Folgen	Ohne Krawatte	35 (80 %)	10 (20 %)	45
	Mit Krawatte	43 (75 %)	14 (25 %)	57
Ohne Folgen	Ohne Krawatte	60 (87 %)	9 (13 %)	69
	Mit Krawatte	54 (88 %)	7 (12 %)	61
Total		192 (83 %)	40 (17 %)	232

Tabelle 17: Rückschaufehler: Resultate der Umfrage 2004

- 666 Überraschend war, dass *weniger* Richterinnen und Richter die Fahrlässigkeit bejahten, wenn die Folgen der Handlung angegeben waren (rund 12 % gegenüber rund 22 % in der Gruppe ohne Folgen; genaue Zahlen siehe Tabelle 17. Die Unterschiede zwischen den Gruppen mit und ohne Folgen sind statistisch signifikant, $\chi^2 = 4,436$; $p = 0,035$). Dies widerspricht der ursprünglichen Hypothese, steht aber möglicherweise im Einklang mit verschiedenen Studien zum Rückschaufehler (siehe Diskussion).

¹⁰⁴⁵ Dies deckt sich mit den Resultaten von STEVEN K. JACOBSON/CHARLES R. BERGER, Communication And Justice: Defendant Attributes And Their Effects on the Severity of His Sentence, Speech Monographs 1974, 282-286, die keinen Schönheits-Hofeffekt feststellten, wenn die Manipulation nur darin bestand, die gleiche Person einmal im Anzug/mit gekämmten Haaren und einmal im T-Shirt/mit ungekämmtem Haar zu präsentieren.

E. Diskussion

- 667 Die Resultate der Umfrage 2003 replizieren qualitativ die Resultate von GUTHRIE/RACHLINSKI/WISTRICH, deren Studie die Vorlage für die Umfrage 2003 bildete. Quantitativ war der Rückschaufehler bei den schweizerischen Richtern weniger stark als bei ihren amerikanischen Kollegen (siehe die Werte von GUTHRIE et. al. in Tabelle 18). Die Bestätigung des erstinstanzlichen Urteils wurde auch von den amerikanischen Richtern der Studie von GUTHRIE et al. insgesamt am häufigsten gewählt. Da Berufungsgerichte sowohl in den USA wie in der Schweiz Ermessensentscheide der Vorinstanz nur zurückhaltend überprüfen, dürfte die vollständige Bestätigung des Urteils in der Praxis am häufigsten sein und der Erfahrung der Richterinnen und Richter entsprechen.

Gruppe ↓	Antworten		
	Geringere Sanktion	Bestätigung	Zurückweisung
Geringere Sanktion	38,6 %	40,4 %	21,1 %
Bestätigung	7,4 %	81,5 %	11,1 %
Zurückweisung	20,4 %	27,8 %	51,9 %

Tabelle 18: Resultate der Studie von GUTHRIE/RACHLINSKI/WISTRICH zum Rückschaufehler¹⁰⁴⁶

- 668 Die Resultate zeigen, dass sich die Richter bei der Beurteilung, welcher Ausgang am wahrscheinlichsten ist, massgeblich vom „tatsächlichen“ Ausgang beeinflussen lassen – ein klassischer Rückschaufehler. Man kann natürlich kritisieren, dass die Richter nur gefragt wurden, welcher Ausgang ihrer Ansicht nach am wahrscheinlichsten war, ohne sie ausdrücklich darauf aufmerksam zu machen, dass sie den bereits bekannten Ausgang bei ihrer Beurteilung ignorieren sollten. Die Fragen sind aber formell identisch.¹⁰⁴⁷ Um zu beurteilen, was am wahrscheinlichsten war, darf man logischerweise nicht darauf abstellen, was tatsächlich geschah. Studien zum Rückschaufehler zeigen zudem, dass die Fragen „was war am wahrscheinlichsten?“, „was hätten Sie vorausgesagt, wenn Sie den Ausgang nicht gekannt hätten?“ und „was hätten andere vorausgesagt, die den Ausgang nicht kannten?“ auch funktional identisch sind und zu den gleichen Ergebnissen führen.¹⁰⁴⁸
- 669 Während die Ergebnisse der Umfrage 2003 den Schluss zulassen, dass die Richterinnen und Richter einem Rückschaufehler unterlagen, widersprechen die Resultate der Umfrage 2004 dieser Hypothese. Die Richter, die die Folgen der Tat kannten, beurteilten die Handlung – resp. die Unterlassung bezüglich Abschränkung und Warnhinweis – eher als pflichtwidrig als die Richter, die die Folgen nicht kannten. Was kann dieses Resultat erklären?
- 670 Zwei Hypothesen drängen sich auf. Die erste ist, dass die Folgen, oder besser gesagt, der Kausalverlauf, der zu den Folgen führte, überraschend schien. Wie die in Fussnoten 1014 und 1015 angeführten Studien belegen, kann ein sehr überraschendes Resultat zur Umkeh-

¹⁰⁴⁶ GUTHRIE/RACHLINSKI/WISTRICH, FN 20, 803.

¹⁰⁴⁷ GUTHRIE/RACHLINSKI/WISTRICH, FN 20, 804.

¹⁰⁴⁸ FISCHHOFF, FN 983, 293 ff.

rung des Rückschaufehlers führen, so dass die Folgen im Nachhinein als völlig unvorhersehbar betrachtet werden. Da keine Daten zum Ausmass der Überraschung erhoben wurden, lässt sich diese These nicht empirisch belegen. Ich weiss nicht, ob die Richter die Folgen der Unterlassung des Bauleiters als überraschender empfanden als den Ausgang des Berufungsverfahrens der ersten Umfrage. Während „Ertrinken im Teich“ kaum eine überraschende Folge einer künstlichen Vertiefung eines natürlichen Teiches ist, ist es das Ertrinken von zwei Kindern gleichzeitig, die zudem im Eis einbrechen – was kaum die mehrheitliche Ursache von Ertrinkungsunfällen ist – möglicherweise schon. Die Hypothese erscheint daher zumindest plausibel, lässt sich aber wie erwähnt aufgrund der erhobenen Daten weder beweisen noch widerlegen.

- 671 Die zweite Hypothese ist, dass die Richter, denen die Folgen der Tat mitgeteilt wurden, ahnten, dass von ihnen erwartet wurde, dass sie die Fahrlässigkeit bejahen und sie diese Erwartung bewusst unterliefen. Ein solches Unterlaufen von vermeintlichen oder tatsächlichen Erwartungen des Versuchsleiters ist bei allzu offensichtlicher Manipulation nicht unbekannt. Auch diese Hypothese ist plausibel, kann aber durch die vorliegenden Daten nicht belegt werden.
- 672 Eine Schwäche der Umfrage 2004 ist, dass praktisch alle Befragten – unabhängig von der Manipulation – die Fahrlässigkeit bejahten. Der Sachverhalt war daher nicht sehr geeignet, Unterschiede in der Beurteilung der Fahrlässigkeit aufzuspüren. Dies lässt sich im Nachhinein nicht mehr ändern. Aufgrund der Resultate der Pilotstudie war ich davon ausgegangen, dass rund ein Drittel der Richterinnen und Richter die Fahrlässigkeit verneinen würden, wie dies die Versuchspersonen des Pretests (Wirtschaftsanwälte einer mittelgrossen Zürcher Kanzlei) getan hatten.

V. Rückschaufehler im Recht: Ignoranz oder Anpassung?

- 673 Dass man im Nachhinein immer schlauer ist, ist eine Volksweisheit. Leuten, die „es schon immer gewusst haben“ begegnen wir in der Regel mit einer gesunden Portion Skepsis. Anders als andere kognitive Täuschungen, derer sich das Rechtssystem und seine Akteure kaum bewusst sind, ist der Rückschaufehler – nicht notwendigerweise die psychologische Forschung dazu – Juristen seit langem bekannt. Entsprechende Warnungen – „*after the event, even a fool is wise*“¹⁰⁴⁹ – finden sich daher auch in der schweizerischen Rechtsprechung.¹⁰⁵⁰ Fraglich ist, ob sie in Anbetracht der psychologischen Literatur, die ihnen jede Wirkung abspricht, sehr nützlich sind. Im Folgenden wird am Beispiel der privatrechtlichen deliktischen Verschuldenshaftung (Art. 41 OR) erörtert, welche der klassischen „Formeln“ des Haftungsrechts Rückschaufehler begünstigen und welche dazu beitragen

¹⁰⁴⁹ Weekly Law Reports 1961/2, 126 ff., 140 – *The Wagon Mound*.

¹⁰⁵⁰ BGE 113 II 429, 432: Es gehe nicht an „aus einer Behandlung oder Operation, die sich nachträglich als unangemessen oder sogar verfehlt erweist, leichthin auf eine haftungsbegründende Vertragsverletzung zu schliessen“; ebenso 130 IV 7, 12. Die in der vorangehenden Fussnote erwähnte „*Wagon Mound*“ Entscheidung wird zitiert von VITO ROBERTO, Schweizerisches Haftpflichtrecht, Zürich 2002, Rz. 191 und 194; PIERRE WIDMER, Privatrechtliche Haftung, in: PETER MÜNCH/THOMAS GEISER (Hrsg.), Schaden – Haftung – Versicherung, Basel etc. 1999, 7-93, Rz. 2.53.

können, sie zu vermeiden. Bei der vertraglichen Haftung und dem strafrechtlichen Fahrlässigkeitsdelikt stellen sich grundsätzlich die gleichen Probleme.

A. Rückschaufehler und privatrechtliche Verschuldenshaftung

- 674 Nach herrschender schweizerischer Lehre und Rechtsprechung sind die vier Voraussetzungen für die ausservertragliche Haftung nach Art. 41 OR Schaden, Rechtswidrigkeit, Kausalzusammenhang und Verschulden. Der Kausalzusammenhang muss sowohl „natürlich“ wie „adäquat“ sein und zwischen der dem Haftenden zurechenbaren Handlung oder Unterlassung und dem Schaden bestehen.¹⁰⁵¹ Die Voraussehbarkeit der Folgen einer Handlung spielt sowohl beim Verschulden als auch beim adäquaten Kausalzusammenhang eine Rolle.

1. Verschulden und Voraussehbarkeit

- 675 „Grundvoraussetzung für das Bestehen einer Sorgfaltspflichtverletzung und mithin für die Fahrlässigkeitshaftung ist die Vorhersehbarkeit des Erfolgs“ (BGE 127 IV 62, 65). „Verschulden setzt voraus, dass der Schädiger die mögliche Verursachung einer Schädigung eines Dritten durch sein Verhalten erkennt oder erkennen kann. [...] Eine nicht voraussehbare Schädigung führt nicht zur Bejahung einer Verschuldenshaftung“.¹⁰⁵² Der Massstab, an dem bei der zivilrechtlichen Haftung die Sorgfaltspflicht, und mithin die Voraussehbarkeit, gemessen wird, ist ein objektiver.¹⁰⁵³ Es kommt mit anderen Worten bei der zivilrechtlichen Haftung – anders als bei der strafrechtlichen Verantwortlichkeit – nicht darauf an, ob der konkrete Schädiger in der Lage war, die Folgen seines Verhaltens vorzusehen, sondern ob ein durchschnittlicher Angehöriger der Gruppe, der der Schädiger angehört, in der Lage gewesen wäre, die schädlichen Folgen seines Verhaltens vorzusehen.¹⁰⁵⁴ Wenn die Lehre daher im Zusammenhang mit der Voraussehbarkeit des Verschuldens von „subjektiver Voraussehbarkeit“ – in Abgrenzung zur „objektiven Voraussehbarkeit“ des adäquaten Kausalzusammenhangs¹⁰⁵⁵ – spricht, dann ist dies zumindest missverständlich. Die Voraussehbarkeit bei der zivilrechtlichen Verschuldenshaftung ist eine objektive, aber sie ist, immerhin, aufgrund der objektiv vorhandenen Kenntnisse in dem Zeitpunkt zu treffen, als der Schädiger die schädigende Handlung vorgenommen hat. Einem Arzt kann beispielsweise nicht zum Vorwurf gereichen, dass er die potenzierende Wirkung von zwei von ihm verabreichten Medikamenten nicht voraussah, wenn dies im Zeitpunkt der Verabreichung erst in engsten Fachkreisen bekannt war (BGE 64 II 202).

¹⁰⁵¹ Statt vieler ROLAND BREHM, Kommentar zu Art. 41-61 OR, 4. Aufl., Bern 1998, Art. 41 Rz. 103 ff.

¹⁰⁵² KARL OFTINGER/EMIL W. STARK, Schweizerisches Haftpflichtrecht, Erster Band: Allgemeiner Teil, 5. Aufl., Zürich 1995, 193.

¹⁰⁵³ Statt vieler OFTINGER/STARK, FN 1052, 205 f.

¹⁰⁵⁴ ROBERTO, FN 1050, 56.

¹⁰⁵⁵ ALFRED KELLER, Haftpflicht im Privatrecht, 6. Aufl., Bern 2002, 86

a. Voraussehbarkeit wird meist bejaht

- 676 Fälle, in denen die Voraussehbarkeit der Folgen eines Verhaltens verneint wurden, sind unter den publizierten Bundesgerichtsentscheiden sowohl in zivil- wie in strafrechtlichen Fällen selten.¹⁰⁵⁶ Typischerweise wird die Voraussehbarkeit dann verneint, wenn es dem Schädiger gelingt nachzuweisen, dass die Fachwelt allgemein (BGE 64 II 202) oder eine anerkannte Fachperson im Zeitpunkt der schädigenden Handlung die Folgen ebenfalls nicht vorausgesehen hat. In BGE 116 IV 182 wurde dem Angeklagten zum Vorwurf gemacht, dass er trotz erkennbarer Lawinengefahr die Strasse zwischen Täsch und Zermatt nicht gesperrt hatte. Das Bundesgericht hielt dafür, dass der Lawinenabgang nicht voraussehbar war. Nicht unerheblich war bei dieser Entscheidung, dass ein Beobachter des Eidgenössischen Instituts für Schnee- und Lawinenforschung in Zermatt in einem Bulletin rund eine Stunde vor dem Lawinenabgang zwar eine grosse und steigende Schneebrettgefahr oberhalb von 2'000 m festgestellt habe, aber keine unmittelbare Lawinengefahr. Das Bulletin war dem Angeklagten zwar nicht bekannt, aber es zeigt, dass die Lawinengefahr auch für einen erfahrenen Fachmann nicht erkennbar war.
- 677 In den Fällen, in denen es an einem *ex ante* Vergleichsstandard fehlt, ist die Gefahr gross, dass im Nachhinein aufgrund des Rückschaufehlers fast alles als voraussehbar betrachtet wird. Wenn während Jahrzehnten die Lawinengefahr immer durch „abtreten“ geprüft wurde, und es nie zu einem Unfall kam, dann ist ein Unfall, der trotz positivem Prüfungsverlauf geschieht, nicht voraussehbar. Die Argumentation des Bundesgerichts, eine Methode könne auch dann fehlerhaft sein, wenn sie üblicherweise angewandt wird und es aus irgendwelchen Gründen während Jahren nie zu Unfällen gekommen sei (BGE 115 IV 189, 196), ist zwar richtig. Aber sie besagt nur, dass die fehlerhafte Methode *in Zukunft* nicht mehr angewandt werden sollte. Durch den Unfall hat sich ergeben, dass die Methode nicht sicher ist. Vor dem Unfall war dies nicht erkennbar, denn wie soll der Handelnde erkennen, dass es nicht seine Methode ist, die Unfälle verhinderte, sondern „irgendwelche anderen Gründe“, wenn diese unbekannten Gründe dazu führten, dass es nie zu einem Unfall kam? Ein Vorwurf – zumal ein strafrechtlicher, wie in dem erwähnten BGE – kann ihm erst gemacht werden, wenn er an seiner mangelhaften Methode trotz des Unfalls auch in Zukunft festhält.¹⁰⁵⁷

b. Die Einhaltung von ex ante formulierten Vorschriften als Exkulpation

- 678 Um den Rückschaufehler zu vermeiden, sollte wenn immer möglich auf die Einhaltung von Regeln abgestellt werden, die im Zeitpunkt der schädigenden Handlung in Kraft waren. Sie umschreiben, wie OFTINGER/STARK zu Recht festhalten, bei welchem Verhalten ein Schaden vorausgesehen werden kann.¹⁰⁵⁸ Die Rechtsprechung anerkennt denn auch die

¹⁰⁵⁶ Fälle in denen eine Haftung resp. ein Verschulden mangels Vorhersehbarkeit verneint wurde, sind BGE 99 IV 127 und 126 III 14.

¹⁰⁵⁷ In BGE 115 IV 189 hätte der Angeklagte die Mangelhaftigkeit seiner Methode erkennen können, weil ihm die Sicherung der Pisten durch Lawinensprengungen bereits in der Betriebskonzession vorgeschrieben worden war. Dies ändert aber nichts daran, dass die Hilfsbegründung des Bundesgerichts unzutreffend ist.

¹⁰⁵⁸ OFTINGER/STARK, FN 1052, 204.

Bedeutung von staatlichen und unter gewissen Umständen auch privaten Vorschriften für die Festlegung des Sorgfaltsmassstabs: „Wo besondere, der Unfallverhütung und der Sicherheit dienende Normen ein bestimmtes Verhalten gebieten, richtet sich das Mass der im Einzelfall zu beachtenden Sorgfalt in erster Linie nach diesen Vorschriften. Das Gleiche gilt für entsprechende allgemein anerkannte Verhaltensregeln, auch wenn diese von einem privaten oder halböffentlichen Verband erlassen wurden und keine Rechtsnormen darstellen“ (BGE 130 IV 7, 11).

- 679 Die Rechtsprechung betont jedoch zugleich, dass die Einhaltung der Regeln nicht in allen Fällen zu exkulpieren vermag: „Das [sc. die Einhaltung einschlägiger Vorschriften] schliesst nicht aus, dass der Vorwurf der Fahrlässigkeit auch auf allgemeine Rechtsgrundsätze wie etwa den allgemeinen Gefahrensatz gestützt werden kann“ (BGE 126 IV 13). Dies wird auch in der Lehre gerne betont.¹⁰⁵⁹
- 680 Überblickt man die Rechtsprechung des Bundesgerichts, kann man sich des Eindrucks nicht erwehren, dass die Verletzung von einschlägigen Normen in erster Linie zur Begründung der Sorgfaltspflichtverletzung verwendet wird, während der Nachweis der Einhaltung der Normen kaum je genügt, sich zu exkulpieren.¹⁰⁶⁰ Entscheide wie BGE 76 II 171, wo ein Verschulden des Landwirts verneint wurde, weil das Gesetz keine Beleuchtung vom Felde heimkehrender landwirtschaftlicher Fuhrwerke vorschrieb, sind meist älter.
- 681 Der Rückschaufehler legt es nahe, dass die Sorgfalt in erster Linie an der Einhaltung einschlägiger Vorschriften gemessen werden sollte. Diese sind eine Kodifikation des Verhaltens, das im Zeitpunkt ihres Erlasses nach menschlichem Ermessen nicht zu Schädigungen Dritter führt. Kommt es trotz Einhaltung der einschlägigen Vorschriften zu Schäden Dritter, deutet dies darauf hin, dass der Schaden nicht voraussehbar war und sollte dazu führen, dass das Verschulden verneint wird.
- 682 Wenn man fordert, dass sich die Sorgfaltswidrigkeit in erster Linie an der Einhaltung von bestehenden Regeln und Vorschriften misst, richtet sich das Augenmerk naturgemäss vermehrt auf diese Vorschriften und ihr Zustand kommen. Zu Recht wird bemerkt, dass die betroffenen Verkehrskreise den Sorgfaltsmassstab nicht selber bestimmen können;¹⁰⁶¹ man darf, bildlich gesprochen, den Bock nicht zum Gärtner machen. Andererseits bleiben staatliche Vorschriften oft hinter dem Stand der Technik zurück.¹⁰⁶²
- 683 Ehe man auf die Vorschriften von privaten oder para-staatlichen Organisationen abstellt, ist daher zu prüfen, wer die Normen erlassen hat. Sind in den entsprechenden Gremien nur Vertreter der potentiell Haftpflichtigen vertreten, oder handelt es sich um paritätisch besetzte Kommissionen? Zweitens, werden die Normen laufend dem Stand der Technik angepasst und, besonders wichtig, gibt es einen Feedback-Mechanismus? Dass man einen Unfall als nicht voraussehbar taxiert, bedeutet nicht, dass man nicht für die Zukunft etwas daraus lernen kann und die Vorschriften anpassen sollte. Solche institutionalisierten Feed-

¹⁰⁵⁹ KELLER, FN 1055, 123; OFTINGER/STARK, FN 1052, 216.

¹⁰⁶⁰ Nachweise von Urteilen, in denen trotz Einhaltung einschlägiger Normen gegen den Schädiger entschieden wurde bei OFTINGER/STARK, FN 1052, 216, Fn. 108.

¹⁰⁶¹ ROBERTO, FN 1050, 18.

¹⁰⁶² ROBERTO, FN 1050, 15.

back-Mechanismen sind daher für die Qualität von Sicherheitsvorschriften zentral; sie bestehen bei der Flugsicherheit und (weniger ausgeprägt) in der Medizin. Die Anpassung der Normen für die Zukunft darf aber nicht, genauso wenig wie das Treffen zusätzlicher Sicherheitsmassnahmen durch den Schädiger nach einem Unfall, als Eingeständnis der Vorausschbarkeit des Schadens gesehen werden.¹⁰⁶³

c. Die Problematik des Gefahrensatzes

- 684 Insbesondere der Gefahrensatz kann im Nachhinein fast jedes Verhalten als sorgfaltswidrig erscheinen lassen.¹⁰⁶⁴ Gemäss dem Gefahrensatz ist die Person, die einen gefährlichen Zustand geschaffen hat oder unterhält, dazu verpflichtet, alle erforderlichen Vorsichtsmassnahmen zu treffen, um zu verhüten, dass sich die Gefahr in einem Schaden verwirklicht.¹⁰⁶⁵ Wie aber erkennt man, ob ein Zustand gefährlich ist? Im Haftpflichtprozess in der Regel daran, dass sich die Gefahr verwirklicht hat. Der Schluss ist dann einfach: aus dem Eintritt der Verletzung wird sowohl auf den gefährlichen Zustand als auch darauf geschlossen, dass nicht alle erforderlichen Massnahmen zu ihrer Verhinderung getroffen wurden.
- 685 Exemplarisch BGE 127 IV 62: Beim Reitunterricht war das von der 10-jährigen Geschädigten gerittene Pferd bereits zwei Mal ausgeschert, als der Reitlehrer befahl, vom Schritt in den Trab zu wechseln. Beide Male liess es sich wieder beruhigen, und das Ausscheren hatte keine weiteren Folgen. Beim dritten Mal jedoch brannte es und mit ihm alle anderen Pferde der Reitgruppe durch. Die Geschädigte stürzte vom Pferd und wurde schwer verletzt. Das Bundesgericht warf dem angeklagten Reitlehrer vor, dass er ein drittes Mal befohlen habe, vom Schritt in den Trab zu wechseln, obwohl das Pferd bereits bei den ersten beiden Malen ausgeschert war. Damit habe er eine Gefahrensituation geschaffen, ohne gleichzeitig die erforderlichen Sicherheitsmassnahmen (wie z.B. das Pferd selber zu reiten) getroffen zu haben und somit seine Sorgfaltspflicht als Reitlehrer verletzt (BGE 127 IV 62, 66 f.).
- 686 Da das Pferd beim Wechsel vom Schritt in Trab durchgebrannt ist, scheint es einleuchtend, dass der Reitlehrer mit seinem Befehl einen gefährlichen Zustand geschaffen hat. Dass das Pferd durchbrennen würde, wissen wir aber erst aus der Rückschau. Die Argumentation, dass zwei Mal nichts geschehen ist und es daher nicht vorausschbar war, dass beim dritten Mal etwas passieren würde, ist aus der prognostischen Perspektive mindestens so überzeugend. Es ist daher fraglich, ob es für den Reitlehrer wirklich erkennbar war, dass er einen gefährlichen Zustand schuf, obwohl sich dieser Schluss natürlich im Nachhinein aufdrängt.

¹⁰⁶³ HEINZ REY, *Ausservertragliches Haftpflichtrecht*, 2. Aufl., Zürich 1998, Rz. 877; KELLER, FN 1055, 122.

¹⁰⁶⁴ WIDMER, FN 1050, 40. WIDMER ordnet den Gefahrensatz der Rechtswidrigkeit zu, während OFTINGER/STARK, FN 1052, 11, KELLER, FN 1055, 120, und das Bundesgericht (BGE 124 III 297, 300) ihn dem Verschulden zuordnen. Ich möchte auf diese Kontroverse hier nicht näher eingehen. Entscheidend ist, wie HEINRICH HONSELL, *Schweizerisches Haftpflichtrecht*, 3. Aufl., Zürich 2000, 45, zu Recht bemerkt, dass mit dem Gefahrensatz Handlungspflichten begründet werden können. Er muss mithin immer dann bemüht werden, wenn dem Schädiger eine Unterlassung vorgeworfen wird – die als solche nie ein absolut geschütztes Recht verletzen kann – und es an einer positiv-rechtlich umschriebenen Handlungspflicht fehlt.

¹⁰⁶⁵ BGE 82 II 28; 93 II 92; 95 II 93, 96; KELLER, FN 1055, 120; WIDMER, FN 1050, 40

2. Adäquater Kausalzusammenhang

- 687 Nach herrschender Lehre und Rechtsprechung muss zwischen dem Verhalten des Haftpflichtigen und dem Schaden sowohl ein natürlicher wie adäquater Kausalzusammenhang bestehen. Der natürliche¹⁰⁶⁶ Kausalzusammenhang zwischen einer Ursache und einem Erfolg besteht, wenn die Ursache nicht weggedacht werden kann, ohne dass der Erfolg entfällt, die Ursache mithin *conditio sine qua non* für den Erfolg ist (BGE 125 IV 195, 197).
- 688 Der Kreis der *conditiones sine quibus non* für einen Erfolg ist praktisch unbegrenzt.¹⁰⁶⁷ Auch die Tatsache, dass der Schädiger geboren wurde, lässt sich nicht wegdenken, ohne dass die Schädigung entfällt. Um den Kreis aller Ursachen auf die rechtlich relevanten Ursachen einzuschränken, wird in der Schweiz die Formel der Adäquanz bemüht. Gemäss der Lehre des adäquaten Kausalzusammenhangs hat „ein Ereignis als adäquate Ursache eines Erfolges zu gelten, wenn es nach dem gewöhnlichen Lauf der Dinge und nach der allgemeinen Lebenserfahrung an sich geeignet ist, einen Erfolg von der Art des eingetretenen herbeizuführen, der Eintritt des Erfolges also durch das Ereignis allgemein als begünstigt erscheint“ (BGE 123 III 110, 112).
- 689 Ob ein Ereignis adäquat kausal ist für einen Erfolg, ist ein Werturteil, und zwar eines, bei dem dem Gericht ein erheblicher Ermessensspielraum zukommt.¹⁰⁶⁸ Nach der neueren Lehre beurteilt sich die Adäquanz *ex post*, d.h. man geht vom eingetretenen Schaden aus und prüft, ob ein bestimmter Umstand als dessen Ursache erscheint. Auf die Erkennbarkeit *ex ante* kommt es dabei nicht an, der Entscheid ergeht aus der Retrospektive.¹⁰⁶⁹
- 690 Bei dieser Sicht der Adäquanz ist es nicht überraschend, dass Urteile, bei denen ein adäquater Kausalzusammenhang verneint wurde, selten sind.¹⁰⁷⁰ Die Rechtsprechung betont denn auch, dass – obwohl ein Ereignis gemäss dem *gewöhnlichen* Lauf der Dinge geeignet sein muss, einen Erfolg von der Art des Eingetretenen zu bewirken – auch „ausserordentliche“, „ungewöhnliche“, „singuläre“ und „seltene“ Folgen adäquat sein können (BGE 119 Ib 334, 345 f.).¹⁰⁷¹
- 691 Für einen Teil der Lehre beweist dies, dass die Adäquanz einem natürlichen Rechtsempfinden entspricht und in Fällen, in denen es an Adäquanz fehlt, selten prozessual Schadenersatz verlangt wird.¹⁰⁷² Andere weisen darauf hin, dass die Lehre von der Adäquanz bei einer Betrachtung *ex post* ihre Funktion der Haftungsbeschränkung weitgehend verliert,

¹⁰⁶⁶ Von „naturwissenschaftlichem“ Kausalzusammenhang sollte man besser nicht sprechen, denn auch der „natürliche“ Kausalzusammenhang beruht auf einem Werturteil; ROBERTO, FN 1050, 45.

¹⁰⁶⁷ WIDMER, FN 1050, 45.

¹⁰⁶⁸ BK-BREHM, Art. 41 N. 163.

¹⁰⁶⁹ BGE 80 II 383, 344; OFTINGER/STARK, FN 1052, 115; REY, FN 1063, Rz. 538.

¹⁰⁷⁰ HONSELL, FN 1064, § 3 N. 10; KELLER, FN 1055, 86; OFTINGER/STARK, FN 1052, 114, Fn. 23; Kasuistik bei BK-BREHM, Art. 41 N. 129.

¹⁰⁷¹ BK-BREHM Art. 41 N. 124; ROBERTO, FN 1050, 53.

¹⁰⁷² OFTINGER/STARK, FN 1052, 114, Fn. 23; ähnlich wohl auch KELLER, FN 1055, 86, der die „Unsitte“ kritisiert, ständig die Adäquanz zu bestreiten.

denn im Nachhinein erscheint kaum je ein Kausalzusammenhang als inadäquat.¹⁰⁷³ KELLER weist darauf hin, dass zumindest bei der Verschuldenshaftung die Prüfung der Adäquanz überflüssig ist, da sie zu keiner stärkeren Beschränkung der Haftung führen kann als die Prüfung der Vorausschbarkeit beim Verschulden. Für ihn hat die Adäquanz in erster Linie bei den Kausalhaftungen ihre Berechtigung als haftungsbeschränkendes Element.¹⁰⁷⁴ Für ROBERTO bleibt für die Haftungsbeschränkung durch die Formel der Adäquanz hingegen bei den Gefährdungshaftungen wenig Raum, denn wenn sich die Gefährdung verwirklicht habe, seien die Folgen immer adäquat, auch wenn sie ungewöhnlich erscheinen mögen.¹⁰⁷⁵

- 692 Tatsächlich scheint die Formel der Adäquanz, wie sie zurzeit in der Schweiz verwendet wird, wenig geeignet, zu einer Beschränkung der Haftung zu führen. Im Nachhinein erscheint praktisch jede Folge einer Ursache als dem gewöhnlichen Lauf der Dinge entsprechend. Es ist schon schwierig genug, in bewusster Missachtung der tatsächlichen Folgen die Wahrscheinlichkeit der möglichen Folgen einer Handlung einzuschätzen. Ist der Ausgang bekannt und berücksichtigt man ihn ganz bewusst, ist es sinnlos, noch von Wahrscheinlichkeit zu sprechen; die Wahrscheinlichkeit, dass sich das Ereignis so abgespielt hat, wie es sich eben abgespielt hat, ist immer 100 %.¹⁰⁷⁶ Die „retrospektive Prognose“¹⁰⁷⁷ ist ein Rezept, den Rückschafehler mit Sicherheit immer zu begehen.
- 693 Will man die Haftung beschränken – letztlich eine rechtspolitische Frage, von ihren Befürwortern wird die Adäquanztheorie auch deshalb verteidigt, weil sie die Haftung nur selten ausschliesst¹⁰⁷⁸ – wird man daher nicht umhin kommen, andere Lehren, sei es die Normzwecklehre oder die Lehre vom „allgemeinen Lebensrisiko“, zu bemühen.¹⁰⁷⁹

¹⁰⁷³ HONSELL, FN 1064, § 3 N. 12 ff.; ROBERTO, FN 1050, 57; WIDMER, FN 1050, 47.

¹⁰⁷⁴ KELLER, FN 1055, 86.

¹⁰⁷⁵ ROBERTO, FN 1050, 57.

¹⁰⁷⁶ ROBERTO, FN 1050, 56 f..

¹⁰⁷⁷ REY, FN 1063, Rz. 539, weist zu Recht darauf hin, dass man bei einer nachträglichen Beurteilung eigentlich nicht mehr von einer Prognose sprechen kann.

¹⁰⁷⁸ BK-BREHM Art. 41 N. 124.

¹⁰⁷⁹ Zur Normzwecklehre und dem „allgemeinen Lebensrisiko“ als Alternativen zur Adäquanz siehe ROBERTO, FN 1050, 59 f.

§ 15 Hofeffekt (*halo effect*)

- 694 Der Hofeffekt (*halo effect*) wurde erstmals 1907 von FREDERIC L. WELLS beobachtet und 1920 von EDWARD L. THORNDIKE benannt.¹⁰⁸⁰ Unter dem Hofeffekt wird die Tendenz verstanden, faktisch unabhängige oder nur mässig korrelierte Eigenschaften von Personen oder Sachen fälschlicherweise als zusammenhängend wahrzunehmen.¹⁰⁸¹ THORNDIKE hatte herausgefunden, dass Instruktionsoffiziere ihre Untergebenen bezüglich Intelligenz, Physis, Führungsverhalten und Charakter oft ähnlich positiv oder negativ beurteilten, obwohl diese Eigenschaften – so THORNDIKE – eigentlich wenig miteinander zu tun haben.¹⁰⁸²
- 695 ASCH bat seine Versuchspersonen, sich einen Eindruck von jemandem zu machen, der als intelligent, fachkundig, fleissig, warm, entschlossen, praktisch und vorsichtig beschrieben wurde, während anderen Versuchspersonen die Person als intelligent, fachkundig, fleissig, kalt, entschlossen, praktisch und vorsichtig vorgestellt wurde. Die beiden Beschreibungen unterscheiden sich also nur dadurch, dass die zu beurteilende Person einmal als „warm“, das andere Mal als „kalt“ beschrieben wird.¹⁰⁸³ Den Versuchspersonen wurden dann Begriffspaare wie grosszügig/geizig, glücklich/unglücklich oder reizbar/freundlich vorgelegt. 75-95 % der Versuchspersonen, denen die Person als „warm“ beschrieben wurde, dachten, dass diese Person auch grosszügig, glücklich, freundlich und humorvoll sei, aber nur 5-35 % der Versuchspersonen kamen zu dieser Einschätzung, wenn die Person als „kalt“ beschrieben wurde.¹⁰⁸⁴ Die positive Eigenschaft „warm“ strahlt auf die Beurteilung weiterer Persönlichkeitseigenschaften hat, sie hat einen Hof (oder wörtlich übersetzt einen Glorien-schein). Der Hofeffekt wird als „allgegenwärtig“ bezeichnet.¹⁰⁸⁵
- 696 Der Hof, der vom Prestige¹⁰⁸⁶ einer Universität ausgeht, führt dazu, dass eine prestigeträchtige Universität mehr Lizenzverträge betreffend ihrer Erfindungen abschliessen kann. Andere Variablen, die einen Einfluss auf die Anzahl abgeschlossener Lizenzverträge haben können, wie die Leistung in der Vergangenheit,¹⁰⁸⁷ Anzahl patentierter Erfindungen, Vorhandensein einer Stelle, die die Erfindungen der Universität vermarktet,

¹⁰⁸⁰ FREDERIC L. WELLS, A Statistical Study of Literary Merit, Archives of Psychology 1907, 1-30; EDWARD L. THORNDIKE, A Constant Error in Psychological Ratings, Journal of Applied Psychology 1920, 25-29, 25.

¹⁰⁸¹ MANFRED SCHMITT, Schönheit und Talent: Untersuchungen zum Verschwinden des Halo-Effekts, Zeitschrift für experimentelle und angewandte Psychologie 1992, 475-492, 493.

¹⁰⁸² THORNDIKE, FN , 1080, 10 ff.

¹⁰⁸³ ASCH, FN 913, 262.

¹⁰⁸⁴ ASCH, FN 913, 263 ff.

¹⁰⁸⁵ WILLIAM H. COOPER, Ubiquitous Halo, Psychological Bulletin 1981, 218-244.

¹⁰⁸⁶ „Prestige“ wurde operationalisiert als kombinierter Rang in den *U.S. News & World Report*, *Gourman Report* und *National Research Council* Ranglisten.

¹⁰⁸⁷ Durchschnittliche Einnahme pro Lizenz, der Anteil Erfindungen, die lizenziert wurden, Anteil Lizenzen, die Einnahmen generieren, Gesamteinkommen aus Lizenzen.

Geldquellen,¹⁰⁸⁸ geographische Lage und das Vorhandensein einer medizinischen Fakultät wurden dabei kontrolliert.¹⁰⁸⁹

- 697 Untersuchungen zum „Schönheits-Hofeffekt“ zeigen, dass schöne Menschen als kompetenter und glücklicher eingeschätzt und ihre Persönlichkeitsmerkmale – nicht nur ihre Schönheit – als erstrebenswerter erachtet werden als diejenigen weniger attraktiver Menschen.¹⁰⁹⁰ LANDY und SIGALL berichten, dass ein Aufsatz, der angeblich von einer schönen Autorin stammt, von männlichen Versuchspersonen besser bewertet wird als der gleiche Aufsatz, der angeblich von einer unattraktiven Autorin stammt.¹⁰⁹¹ Die Resultate von LANDY/SIGALL konnten von SCHMITT mit deutschen Versuchspersonen allerdings nicht repliziert werden.¹⁰⁹² Der Schönheits-Hofeffekt gilt allgemein als robust; FEINGOLD kommt in einer Analyse von 30 Studien zum Schönheits-Hofeffekt zum Schluss, dass schöne Menschen ungeachtet der Versuchspersonen und Stimulus-Materialien (Foto, Live, Video) als sozialkompetenter, dominanter, geistig gesünder, wärmer und intelligenter als unattraktive Menschen betrachtet werden, es aber nicht sind.¹⁰⁹³
- 698 Lange wurde angenommen, dass der Schönheits-Hofeffekt unbewusst wirkt,¹⁰⁹⁴ und dass es nichts nützt, die Probanden explizit darauf hinzuweisen, ihn zu meiden.¹⁰⁹⁵ Neuere Studien zeigen, dass sich die Menschen der Gefahr der Beeinflussung durch die physische Attraktivität eines Mitmenschen durchaus bewusst sind und erfolgreich – oder besser gesagt: ins andere Extrem verfallend – darauf reagieren können.¹⁰⁹⁶ Gegen die Annahme, dass der Hofeffekt gänzlich unbewusst wirkt, spricht auch, dass er bei Zeitdruck stärker ist; haben die Versuchspersonen m. a. W. mehr Zeit, so können sie dem Einfluss des Hofeffekts zumindest teilweise entgegenwirken.¹⁰⁹⁷ SCHMITT fand einen Hofeffekt, wenn nur die

¹⁰⁸⁸ Universitäten, die stärker privat finanziert sind, produzieren vermutlich mehr kommerziell verwertbare Erfindungen als primär staatlich finanzierte Universitäten.

¹⁰⁸⁹ WESLEY DAVID SINE/SCOTT SHANE/DANTE DI GREGORIO, The Halo Effect and Technology Licensing: The Influence of Institutional Prestige on the Licensing of University Inventions, *Management Science* 2003, 478-496.

¹⁰⁹⁰ KAREN DION/ELLEN BERSCHIED/ELAINE WALSTER, What Is Beautiful Is Good, *Journal of Personality and Social Psychology* 1972, 285-290, 288.

¹⁰⁹¹ DAVID LANDY/HAROLD SIGALL, Beauty Is Talent: Task Evaluation as a Function of the Performer's Physical Attractiveness, *Journal of Personality & Social Psychology* 1974, 299-304; ähnlich MARCIA J. MURPHY/DAVID T. HELKAMP, Attractiveness and Personality Warmth: Evaluations of Paintings Rated by College Men and Women, *Perceptual & Motor Skills* 1976, 1163-1166: schönere Menschen werden als talentiertere Maler empfunden.

¹⁰⁹² SCHMITT, FN 1081, 479.

¹⁰⁹³ ALAN FEINGOLD, Good-looking people are not what we think, *Psychological Bulletin* 1992, 304-341. Hingegen argumentieren SATOSHI KANAZAWA/JODY L. KOVAR, Why Beautiful People Are More Intelligent, *Intelligence* 2004, 227-243, dass schöne Menschen intelligenter sind, weil (a) sie oft einen höheren sozialen Rang einnehmen; (b) höherrangige Männer im Schnitt mit schöneren Frauen Kinder zeugen; (c) Intelligenz und (d) Schönheit vererblich sind. KANAZAWA/KOVAR belegen ihre vier Prämissen mit empirischen Daten.

¹⁰⁹⁴ SCHMITT, FN 1081, 477.

¹⁰⁹⁵ CHRISTOPHER G. WETZEL/TIMOTHY D. WILSON/JAMES KORT, The Halo Effect Revisited: Forewarned Is Not Forearmed, *Journal of Experimental Social Psychology* 1981, 427-439.

¹⁰⁹⁶ RONALD M. FRIEND/MICHAEL VINSON, Leaning Over Backwards: Jurors' Responses to Defendants' Attractiveness, *Journal of Communication* 1974, 124-129.

¹⁰⁹⁷ KARL CHRISTOPH KLAUER/ARMIN SCHMELING, Sind Halo-Fehler Flüchtigkeitsfehler?, *Zeitschrift für experimentelle und angewandte Psychologie* 1990, 594-607.

Leistung, nicht aber die Persönlichkeit der zu beurteilenden Person bewertet wurde. Wurde hingegen zuerst die Persönlichkeit bewertet, zeigte die anschliessende Leistungsbeurteilung keinen Einfluss des Hofeffekts. SCHMITT schreibt dies einem Wiedergutmachungsmotiv der Versuchspersonen zu, was dadurch gestützt wird, dass die Leistungen von Personen, denen unverschuldetes Unglück zustiess, besser bewertet wurden.¹⁰⁹⁸

- 699 Der Schönheits-Hofeffekt wirkt gemäss MARIO A. BENASSI nur dann, wenn der Stimulus – auch hier ein Foto – den Versuchspersonen *vor* der Leistung der zu beurteilenden Person präsentiert wird. Erhalten die Versuchspersonen hingegen zuerst Informationen über die Leistung der zu beurteilenden Person und erst dann Informationen zu ihrer physischen Attraktivität, so zeigt sich kein Hofeffekt.¹⁰⁹⁹ Da der Beurteilende in den meisten Fällen die Attraktivität einer Person vor ihrer Leistung zur Kenntnis nimmt – man denke nur an typische Beurteilungssituationen im Arbeitsleben oder in der Schule – hat dies möglicherweise geringe Auswirkungen auf viele Beurteilungssituationen. Aufgrund des schweizerischen Strafprozessrechts wird sich das Gericht jedoch meist bereits schon vor der Hauptverhandlung – wo es den Angeklagten erstmals sieht, da meist keine Fotos des Angeklagten in den Untersuchungsakten sind – ein Bild über die „Leistung“ des Angeklagten gemacht haben, so dass der Befund von BENASSI darauf hindeutet, dass der Schönheits-Hofeffekt möglicherweise in der gerichtlichen Praxis der Schweiz geringere Bedeutung hat als die weiter hinten vorgestellten ausländischen Studien suggerieren.
- 700 Als Ursache für den Hofeffekt kommen in Frage: (i) der Einfluss des generellen Eindrucks auf die Beurteilung einzelner Persönlichkeitsmerkmale („*General Impression Model*“); (ii) der Einfluss einer herausragenden Eigenschaft auf die Beurteilung anderer Eigenschaften der gleichen Person („*Salient Dimension Model*“); oder (iii) die ungenügende Unterscheidung konzeptionell verschiedener Eigenschaften der beurteilten Person („*Inadequate Discrimination Model*“). In vielen praktischen Fällen lassen sich zumindest das „General Impression“ und das „Salient Dimension“ Modell kaum unterscheiden; LANCE et. al. kommen jedoch aufgrund einer Serie intelligent geplanter Experimente zum Schluss, dass das Modell des generellen Eindrucks den Hofeffekt am besten und unter verschiedensten Bedingungen zu erklären vermag.¹¹⁰⁰

I. Kritik an der Forschung zum Hofeffekt

- 701 MURPHY, JAKO und ANHALT haben in einem 1993 erschienen Artikel die Forschung zum Hofeffekt scharf kritisiert.¹¹⁰¹ Ihre Hauptkritikpunkte sind erstens, dass der Hofeffekt nur dort auftrete, wo die Eigenschaften tatsächlich gering korrelierten; zweitens, dass die *beobachtete* Korrelation oft als Mass für den Hofeffekt genommen werde; und drittens,

¹⁰⁹⁸ SCHMITT, FN 1081, 488 ff. Die Resultate von SCHMITT sind allerdings teilweise nicht signifikant.

¹⁰⁹⁹ MARIO A. BENASSI, Effects of Order of Presentation, Primacy, and Physical Attractiveness on Attribution of Ability, *Journal of Personality and Social Psychology* 1982, 48-58.

¹¹⁰⁰ CHARLES E. LANCE/JULIE A. LAPOINTE/AMY M. STEWART, A Test of the Context Dependency of Three Causal Models of Halo Rater Error, *Journal of Applied Psychology* 1994, 332-340.

¹¹⁰¹ KEVIN R. MURPHY/ROBERT A. JAKO/REBECCA L. ANHALT, Nature and Consequences of Halo Error: A Critical Analysis, *Journal of Applied Psychology* 1993, 218-225.

dass der Hofeffekt als schädlicher Fehler betrachtet werde, obwohl er im wirklichen Leben oft gar keine schädlichen Auswirkungen habe.

- 702 Nach dem theoretischen Modell des Hofeffekts müsste er auch dort auftreten, wo die tatsächliche Korrelation der beurteilten Persönlichkeitsmerkmale hoch ist. Die beobachtete Korrelation müsste, wenn der generelle Eindruck zum Hofeffekt führt, noch höher sein als die tatsächliche.¹¹⁰² MURPHY et. al führen an, es gebe keine einzige Studie, die bei hoher tatsächlicher Korrelation der beurteilten Eigenschaften einen Hofeffekt gefunden habe.¹¹⁰³
- 703 Um von einem Hofeffekt sprechen zu können, muss man die tatsächliche Korrelation der Persönlichkeitsmerkmale kennen, ehe man von der beobachteten (d.h. von den Versuchspersonen attribuierten) Korrelation auf einen Fehler schliessen kann. Wenn „warme“ Menschen in der Regel auch „freundlich“ sind, dann ist es kein Fehler der Versuchspersonen von ASCH, wenn sie annehmen, dass ein als „warm“ beschriebener Mensch auch freundlich ist. Während sich in experimentellen Versuchen tatsächliche und illusorische Korrelation unterscheiden lassen – indem man beispielsweise eine Kontrollgruppe die gleiche Person beurteilen lässt, die weder als „warm“ noch als „kalt“ beschrieben wurde – lässt sich nach MURPHY et al. in Wirklichkeit tatsächliche und illusorische Korrelation aus methodischen Gründen nie trennen.¹¹⁰⁴ Schliesslich weisen sie darauf hin, dass es in Wirklichkeit oft nicht sehr wichtig ist, zwischen den einzelnen Persönlichkeitsmerkmalen einer Person zu unterscheiden, sondern zwischen den Persönlichkeiten verschiedener Personen. Ob jemand befördert wird, hängt davon ab, an welcher Stelle in einer Liste aller Kandidaten er steht (*rank order*). Gemäss MURPHY et al. nimmt die Verlässlichkeit des Vergleichs verschiedener Personen zu, wenn die Korrelation der einzelnen Persönlichkeitsmerkmale jeder einzelnen Person grösser ist, so dass der Hofeffekt zu *besseren* Ergebnissen führen kann.¹¹⁰⁵
- 704 MURPHY et al. mögen Recht haben, dass der Hofeffekt bei vielen Beurteilungen keine schädlichen Auswirkungen hat. Für den Schönheits-Hofeffekt bei der Strafzumessung gilt dies aus verschiedenen Gründen nicht: erstens ist die Korrelation von physischer Attraktivität und strafmindernden Umständen mit an Sicherheit grenzender Wahrscheinlichkeit gering, so dass der erste Einwand, dass der Hofeffekt bei hoher tatsächlicher Korrelation verschwindet, nicht greift. Zweitens spielt die Unterscheidung der verschiedenen Ursachen, die zu einer Straftat geführt haben, bei der Strafzumessung eine Rolle. Anders als bei Beförderungsentscheidungen ist nicht die inter-individuelle Differenzierung (*rank order*) wichtig, sondern die Gründe, die im zu beurteilenden Einzelfall für eine hohe oder tiefe Strafe sprechen.

¹¹⁰² MURPHY/JAKO/ANHALT, FN 1101, 219. Die Stärke des Hofeffekts, definiert als Differenz von beobachteter und tatsächlicher Korrelation, ist natürlich geringer, wenn die tatsächliche Korrelation grösser ist.

¹¹⁰³ MURPHY/JAKO/ANHALT, FN 1101, 219.

¹¹⁰⁴ MURPHY/JAKO/ANHALT, FN 1101, 220 f.

¹¹⁰⁵ MURPHY/JAKO/ANHALT, FN 1101, 219 f.

II. Hofeffekt und Strafzumessung

A. Theoretisches Modell

- 705 Nach dem Gesetz bemisst sich die Strafe nach dem Verschulden des Täters, wobei der Richter die Beweggründe, das Vorleben und die persönlichen Verhältnisse des Täters berücksichtigt (Art. 63 StGB). Der Richter macht sich ein umfassendes Bild vom Täter und den Umständen, die zur Tat geführt haben. Situative, oder externe, Faktoren, die die Tat erklären, wirken dabei schuld mindernd, während interne (der Veranlagung oder dem Willen des Täters entsprechende) Faktoren strafferhöhend wirken. Das leuchtet unmittelbar ein, wenn man sich den Extremfall eines externen Faktors, die *vis absoluta*, vorstellt: die Tat wird in diesem Fall ausschliesslich durch externe Faktoren bestimmt, was dazu führt, dass sie dem Täter nicht mehr zugerechnet werden kann. Der (praktisch seltene) Fall der *vis absoluta* führt daher mangels subjektiver Vorwerfbarkeit der Tat zum Freispruch (BGE 104 IV 186, 189).
- 706 Dass die strafmindernde Wirkung externer Faktoren dem Willen des Gesetzgebers entspricht, sieht man leicht, wenn man sich die Strafmilderungsgründe des Art. 64 StGB vor Augen führt: der Richter kann die Strafe mildern, wenn der Täter gehandelt hat (i) in schwerer Bedrängnis, (ii) unter dem Eindruck einer schweren Drohung, (iii) auf Veranlassung einer Person, der er Gehorsam schuldig oder von der er abhängig ist; (iv) wenn der Täter durch das Verhalten des Verletzten ernstlich in Versuchung geführt wurde oder (v) wenn Zorn oder grosser Schmerz über eine ungerechte Reizung oder Kränkung ihn hingeworfen hat. Erreicht die Intensität der Einwirkung auf den Willen des Täters – beispielsweise die Drohung – nicht das für eine Strafmilderung erforderliche Ausmass, darf der Richter sie immer noch strafmindernd (unter Beibehaltung des ordentlichen Strafrahmens) berücksichtigen.¹¹⁰⁶
- 707 Die Strafzumessung lässt sich daher in zwei Stufen unterteilen:¹¹⁰⁷ in einem ersten Schritt wägt der Richter das Verhältnis interner und externer Faktoren ab, die zur Tat geführt haben, in einem zweiten Schritt fällt er aufgrund dieser Abwägung die Strafe aus. Der Einfluss des Hofeffekts bewirkt, dass die Tat eines Täters, von dem der Richter einen positiven ersten Eindruck gewonnen hat, eher auf externe Faktoren zurückgeführt wird, was wiederum zu einer geringeren Strafe führt.
- 708 Experimentelle Befunde bestätigen sowohl, dass eine interne Attribution von Verbrechenursachen zu härteren Strafen führt,¹¹⁰⁸ als auch, dass einem attraktiven Täter eher zu Gute

¹¹⁰⁶ HANS WIPRÄCHTIGER, in: NIGGLI/WIPRÄCHTIGER, FN 959, Art. 64 N 2.

¹¹⁰⁷ HANS WERNER BIERHOFF/ERNST BUCK/RENATE KLEIN, Attractiveness and Respectability of the Offender as Factors in the Evaluation of Criminal Cases, in: HERMANN WEGENER/FRIEDRICH LÖSEL/JOCHEN HAISCH (Hrsg.), Criminal Behavior and the Justice System: Psychological Perspectives, New York etc. 1989, 193-207, 194.

¹¹⁰⁸ JOHN S. CARROLL/JOHN W. PAYNE, Crime Seriousness, Recidivism Risk, and Causal Attributions in Judgments of Prison Term by Students and Experts, Journal of Applied Social Psychology 1977, 595-602 (signifikant war der Einfluss allerdings nur für Studierende, nicht bei den Experten [*parole board members*], die sich offenbar weigerten, aufgrund der kurzen Beschreibungen eine Strafe zu verhängen); JOCHEN HAISCH, Anwendung von Attributionstheorie als normatives Modell für eine rationale Strafzumessung: Experimentelle Überprüfung eines Trainingsprogrammes, Zeitschrift für angewandte Psychologie 1980, 415-234

gehalten wird, dass externe Faktoren seine Tat (mit-)verursacht haben.¹¹⁰⁹ BIERHOFF et al. berichten, dass ihre Versuchspersonen die Gründe für einen Autounfall eher in externen Faktoren (wie der schlecht beleuchteten Strasse) als in internen Faktoren (wie fehlender Aufmerksamkeit des Fahrers) suchten, wenn der Fahrer als „freundlich und grosszügig“ beschrieben wurde, hingegen vermehrt interne Faktoren bemühten, wenn der Fahrer als „Störenfried“ (*troublemaker*) beschrieben wurde.¹¹¹⁰ Plakativ gesagt nimmt der Urteilende an, dass ein „prinzipiell guter“ Mensch aufgrund unglücklicher Umstände straffällig wurde, während ein „prinzipiell böser“ Mensch seinem bösen Willen entsprechend gehandelt hat.

- 709 Im Folgenden werden ausschliesslich Studien referiert, die den Einfluss physischer Attraktivität des Täters auf die Wahrscheinlichkeit der Verurteilung und das Strafmass untersuchen. Zwei Gründe sprechen dafür: erstens wird, besonders in der englischsprachigen Literatur, unter dem Begriff „Attraktivität“ sehr vieles verstanden – die familiäre Position des Angeklagten, seine berufliche Stellung, sein Vorleben oder seine Persönlichkeit.¹¹¹¹ Nicht jede Korrelation der so weit verstandenen (Un-)Attraktivität des Angeklagten mit dem Strafmass muss normativ ein Fehler sein. Das Vorleben des Angeklagten, das ihn unattraktiv erscheinen lässt, darf durchaus zu einer härteren Strafe führen.¹¹¹² Andererseits kann der Befund, dass ein Bankdirektor und ein Angestellter des Putzpersonals der Bank für die gleiche Tat mit dem gleichen Strafmass bestraft werden, nicht dahingehend verstanden werden, dass die Strafe dieselbe ist: es lässt sich zumindest plausibel argumentieren, dass die Strafempfindlichkeit des Bankdirektors höher ist, so dass ihn die nominal gleiche Strafe härter trifft.
- 710 Zweitens weiss man nicht, wie hoch die *tatsächliche* Korrelation zwischen „psychologischer Attraktivität“ – meist ausgedrückt in Begriffen wie freundlich, warm, offen, herzlich – und dem Empfinden von Reue über die Tat ist. Ein Täter, der seine Tat aufrichtig bereut, darf unstrittig geringer bestraft werden als ein Täter, der die gleiche Tat nicht bereut.¹¹¹³ Wenn nun die Versuchspersonen automatisch annehmen, dass ein warmer, freundlicher Angeklagter seine Tat mehr bereut als ein unfreundlicher, kalter, so muss das kein Fehler sein. Solange nicht bekannt ist, ob tatsächlich ein Zusammenhang zwischen Warmherzigkeit und Reue besteht, was zumindest nicht ausgeschlossen erscheint, sollte man den Schluss nicht als Fehler bezeichnen.

427, 423; BIERHOFF/BUCK/KLEIN, FN 1107, 203; siehe aber JOCHEN HAISCH, Verhaltensursachen bei Straftätern und Strafurteile durch Juristen und Laien, Monatsschrift für Kriminologie 1977, 82-88, der keinen Einfluss der internen Attribution auf das Strafmass feststellt.

¹¹⁰⁹ ROBERT MCFATTER, Sentencing Strategies and Justice: Effects of Punishment Philosophy On Sentencing Decisions, Journal of Personality and Social Psychology 1978, 1490-1500.

¹¹¹⁰ BIERHOFF/BUCK/KLEIN, FN 1107, 199.

¹¹¹¹ Übersicht bei ELWORK/SALES/SUGGS, FN 377, 25 f.; FRANCIS C. DANE/LAWRENCE S. WRIGHTSMAN, Effects of Defendants' and Victims' Characteristics on Jurors' Verdicts, in: NORBERT L. KERR/ROBERT M. BRAY (Hrsg.), The Psychology of the Courtroom, New York 1982, 83-115, 101 ff.

¹¹¹² Ebenso WAYNE WEITEN, The Attraction-Leniency Effect in Jury Research: An Examination of External Validity, Journal of Applied Social Psychology 1980, 340-347, 345.

¹¹¹³ WIPRÄCHTIGER, FN 1106, Art. 63 N 111.

- 711 Bei der physischen Attraktivität scheinen mir diese Probleme nicht vorhanden: es ist, wie bereits erwähnt, sehr unwahrscheinlich, dass bei schönen Menschen vermehrt Strafminderungsgründe vorliegen oder dass ihre Strafempfindlichkeit generell höher ist als diejenige unattraktiver Menschen. Eine durchwegs mildere Bestrafung schöner Menschen kann man daher gefahrlos als Fehler, ja als ungerecht bezeichnen. Justitia trägt nicht ohne Grund eine Augenbinde.

B. Experimente

- 712 EFRAN berichtet, dass seine Versuchspersonen – von denen zuvor 93 % gesagt hatten, dass Schönheit bei der Strafzumessung keine Rolle spielen sollte – sich über die Schuld physisch attraktiver Angeklagter weniger sicher waren und die attraktiven Angeklagten weniger hart bestraften. Der Effekt war bei Männern ausgeprägter als bei Frauen.¹¹¹⁴ LEVENTHAL und KRATE berichten ebenfalls, dass attraktivere Angeklagte zu geringeren Strafen verurteilt wurden.¹¹¹⁵
- 713 Nach DARBY und JEFFERS werden attraktive Angeklagte weniger oft für schuldig befunden und erhalten weniger harte Strafen. Der Effekt ist stärker, wenn die Urteilenden sich selbst als attraktiv einschätzen. DARBY/JEFFERS nehmen an, dass Urteilende, die sich selbst als unattraktiv einschätzen, eher mit den unattraktiven Angeklagten sympathisieren und diese milder beurteilen.¹¹¹⁶
- 714 Gemäss JOCHEN PIEHL wird die attraktive Verursacherin eines Verkehrsunfalls weniger hart bestraft, aber dieser Effekt kehrt sich um, wenn der Unfall tödliche Folgen hat.¹¹¹⁷ Einen weiteren Wechselwirkungseffekt zwischen der Attraktivität des Angeklagten und der Art der ihm vorgeworfenen Tat haben HAROLD SIGALL und NANCY OSTROVE in einer viel zitierten Studie gefunden: wenn die Tat nichts mit der Attraktivität der Angeklagten zu tun hat (z.B. Einbruchdiebstahl), wird die attraktive Angeklagte weniger hart bestraft, benutzt sie aber ihre Attraktivität für die Straftat (z.B. Betrug, wobei im konkreten Fall eine attraktive Frau das Vertrauen eines Mannes missbraucht), so wird sie für die gleiche Tat härter bestraft als die unattraktive Angeklagte.¹¹¹⁸ SIGALL/OSTROVE erklären dies damit, dass ein attraktiverer Mensch mehr Gelegenheit für Hochstapelei hat und eher wieder delinquent als ein unattraktiver Betrüger. Die Versuchspersonen, die dies antizi-

¹¹¹⁴ MICHAEL G. EFRAN, The Effect of Physical Appearance on the Judgment of Guilt, Interpersonal Attraction, and Severity of Recommended Punishment in a Simulated Jury Task, *Journal of Research in Personality* 1974, 45-54.

¹¹¹⁵ GLORIA LEVENTHAL/RONALD KRATE, Physical Attractiveness and Severity of Sentencing, *Psychological Reports* 1977, 315-318.

¹¹¹⁶ BRUCE W. DARBY/DEVON JEFFERS, The Effects of Defendant and Juror Attractiveness on Simulated Courtroom Trial Decisions, *Social Behavior and Personality* 1988, 39-50.

¹¹¹⁷ JOCHEN PIEHL, Integration of Information in the "Courts": Influence of Physical Attractiveness on Amount of Punishment for a Traffic Offender, *Psychological Reports* 1977, 551-556. RICHARD R. IZZETT/WALTER LEGINSKI, Group Discussion and the Influence of Defendant Characteristics in a Simulated Jury Setting 1974, 271-279, 275, berichten aber auch bei fahrlässiger Tötung von einem signifikanten Einfluss der Attraktivität der Angeklagten auf das Strafmaß.

¹¹¹⁸ HAROLD SIGALL/NANCY OSTROVE, Beautiful but Dangerous: Effects of Offender Attractiveness and Nature of Crime on Juridic Judgment, *Journal of Personality and Social Psychology* 1975, 410-414.

pieren, bestrafen ihn deswegen aus spezialpräventiven Überlegungen härter.¹¹¹⁹ Die härtere Bestrafung im Fall, in dem die Attraktivität als Vorteil erscheint, kann auch durch das attributionstheoretische Modell erklärt werden: wenn ein Angeklagter sich seine Attraktivität bewusst zu Nutze macht, ist das ein interner Faktor, der zu Straferhöhung führt.¹¹²⁰

- 715 Die Resultate von SIGALL/OSTROVE konnten teilweise repliziert werden von SMITH/HED; der attraktivere Angeklagte wurde für Einbruchdiebstahl mit durchschnittlich 3,5 Jahren Freiheitsstrafe, der unattraktive Angeklagte aber mit 5,7 Jahren bestraft. Bei Hochstapelei fanden SMITH/HED keine Umkehr, aber der Unterschied in der Bestrafung war nicht mehr signifikant.¹¹²¹ Die Resultate von SIGALL/OSTROVE konnten jedoch von WUENSCH et al. in einer 1991 publizierten Studie nicht repliziert werden: WUENSCH et al. berichten, dass der attraktive Angeklagte auch bei Hochstapelei signifikant weniger hart bestraft wird. Frauen waren stärker von der Attraktivität des Angeklagten beeinflusst als Männer (entgegen EFRAN), wobei eine Wirkung immer nur dann vorhanden war, wenn Angeklagte des anderen Geschlechts beurteilt wurden.¹¹²²
- 716 FRIEND und VINSON berichten, dass sich der Effekt in sein Gegenteil verkehrt, wenn die Versuchspersonen explizit gebeten werden, unparteiisch zu sein und ihren persönlichen Eindruck der Attraktivität des Angeklagten nicht das Strafmass beeinflussen zu lassen. Während die Versuchspersonen ohne Ermahnung die attraktive Angeklagte weniger hart bestraften (fünf verglichen mit acht Jahren für die unattraktive Angeklagte), bestraften sie nach der Ermahnung die attraktive Angeklagte mit 8,4 Jahren, die unattraktive aber mit nur 5,5 Jahren Freiheitsstrafe.¹¹²³
- 717 Unklar ist, ob die Beratung in einer Gruppe dazu führt, dass die Bevorzugung attraktiver Angeklagter verschwindet. Gemäss IZZETT und LEGINSKI führt die Beratung in der Gruppe dazu, dass die *sozial und moralisch* unattraktiven Angeklagten weniger hart bestraft werden, während sich die Bestrafung der sozial und moralisch attraktiven Angeklagten nicht signifikant ändert.¹¹²⁴ Interessanterweise führte jedoch die Gruppendiskussion bei SMITH/HED nicht dazu, dass der Effekt der *physischen* Attraktivität verschwand.¹¹²⁵ Dies ist möglicherweise darauf zurückzuführen, dass der Effekt der physischen Attraktivität schwerer zu erkennen ist und daher nicht Thema der Diskussion wird.
- 718 BAUMEISTER und DARLEY berichten, dass der Einfluss sozialer und moralischer Attraktivität auf die Strafzumessung abnimmt, wenn die Urteilenden mehr relevante Fakten zum Verhalten des Angeklagten kennen.¹¹²⁶ Die soziale und moralische Attraktivität

¹¹¹⁹ SIGALL/OSTROVE, FN 1118, 411.

¹¹²⁰ BIERHOFF/BUCK/KLEIN, FN 1107, 197.

¹¹²¹ EDWARD D. SMITH/ANITA HED, Effects of Offender Age and Attractiveness on Sentencing by Mock Jurors, *Psychological Reports* 1979, 691-694.

¹¹²² KARL L. WUENSCH/WILBUR A. CASELOW/CHARLES H. MOORE, Effects of Defendant Attractiveness and Type of Crime on Juridic Judgment, *Journal of Social Behavior and Personality* 1991, 713-724

¹¹²³ FRIEND/VINSON, FN 1096, 127.

¹¹²⁴ IZZETT/LEGINSKI, FN 1117, 277.

¹¹²⁵ SMITH/HED, FN 1121, 692.

¹¹²⁶ ROY F. BAUMEISTER/JOHN F. DARLEY, Reducing the Biasing Effect of Perpetrator Attractiveness in Jury Simulation, *Personality and Social Psychology Bulletin* 1982, 286-292.

dient nach ihren Erkenntnissen vor allem dann als Faktor der Strafzumessung, wenn keine geeigneteren Angaben vorhanden sind. Wenn dies auch für die physische Attraktivität gilt, ist die Bedeutung des Attraktivitäts-Effekts in Gerichtsverfahren möglicherweise gering, da dort zusätzliche Informationen zum Verhalten des Angeklagten vorhanden sind.

- 719 Während die bisher vorgestellten Studien alle kurze schriftliche Sachverhaltsschilderungen und Fotografien der Angeklagten als Stimulus-Material verwendeten, war das Stimulus-Material von KULKA/KESSLER aufwendiger. In ihrer Studie hörten die Versuchspersonen ein Tonband der Plädoyers der Anwälte und der Aussagen des medizinischen Sachverständigen in einem Zivilprozess, in dem es um die Haftung für einen Autounfall ging. Das Tonband wurde von Dias der jeweils sprechenden Person begleitet. War der Kläger attraktiv und der Beklagte unattraktiv, wurden im Schnitt \$ 10'051 Schadenersatz zugesprochen, war der Kläger dagegen unattraktiv und der Beklagte attraktiv, nur \$ 5'623.¹¹²⁷
- 720 Nicht alle Experimente haben jedoch einen Einfluss physischer Attraktivität auf das Strafmass oder die Verurteilungswahrscheinlichkeit gefunden. Keinen Effekt haben beispielsweise JACOBSON und BERGER festgestellt, wenn die Manipulation nur darin bestand, die gleiche Person einmal im Anzug/mit gekämmten Haaren und einmal im T-Shirt/mit ungekämmtem Haar zu präsentieren.¹¹²⁸ Ebenfalls keinen Einfluss der physischen Attraktivität auf Verschulden fanden BURKE et al., wenn es um einen ambivalenten Fall häuslicher Gewalt ging.¹¹²⁹ Nach WEITEN genügt eine Instruktion über die massgeblichen rechtlichen Regeln, wie sie in Geschworenengerichtsverfahren üblich ist, um den Einfluss der sozialen Attraktivität auf die Wahrscheinlichkeit der Verurteilung (nicht aber auf das Strafmass)¹¹³⁰ zum Verschwinden zu bringen.¹¹³¹ Es ist daher besonders interessant zu untersuchen, ob attraktive Angeklagte auch in wirklichen Gerichtsfällen milder beurteilt werden.

C. Feldstudien

- 721 Die zwei mir bekannten Feldstudien, die zu diesem Thema durchgeführt wurden, kommen beide zum Schluss, dass die physische Attraktivität des Angeklagten, resp. des Klägers, einen Einfluss auf die Strafe resp. den Erfolg der Klage hat.
- 722 JOHN E. STEWART liess insgesamt 74 Angeklagte (bis auf vier alle männlich) in ebenso vielen Strafverfahren vor Bezirksgerichten in Pennsylvania von zehn Beobachtern observie-

¹¹²⁷ RICHARD A. KULKA/JOAN B. KESSLER, Is Justice Really Blind? – The Influence of Litigant Physical Attractiveness on Juridical Judgment, *Journal of Applied Social Psychology* 1978, 366-381, 373. Die Kontrollgruppe (keine Dias) sprach im Schnitt \$ 8'618 zu. Ähnlich auch COOKIE STEPHAN/JUDY CORDER TULLY, The Influence of Physical Attractiveness of a Plaintiff on the Decisions of Simulated Jurors, *Journal of Social Psychology* 1977, 149-150.

¹¹²⁸ JACOBSON/BERGER, FN 1045, 286.

¹¹²⁹ DONALD M. BURKE/M. ASHLEY AMES/ROSALIE ETHERINGTON/JEAN PIETSCH, Effects of Victim's and Defendants Physical Attractiveness on the Perception of Responsibility in an Ambiguous Domestic Violence Case, *Journal of Family Violence* 1990, 199-207.

¹¹³⁰ Der Einfluss der *sozialen* Attraktivität auf das Strafmass kann nach WEITEN und der hier vertretenen Auffassung nicht als Fehler bezeichnet werden, siehe Rz. 709.

¹¹³¹ WEITEN, FN 1112, 345.

ren. Die Beobachter mussten verschiedene Eigenschaften der Angeklagten in einem Fragebogen festhalten, eine davon war die physische Attraktivität, wobei die Beobachter nicht wussten, dass dies die entscheidende unabhängige Variable war.¹¹³² Wurden andere Variablen wie Hautfarbe und Schwere der Tat kontrolliert, zeigte sich ein signifikanter, wenn auch schwacher, Einfluss der physischen Attraktivität auf das Strafmass, nicht aber auf die Wahrscheinlichkeit der Verurteilung.¹¹³³

- 723 LESLIE A. ZEBROWITZ und SUSAN M. McDONALD liessen 506 Fälle beobachten, die von *Small Claims Courts* in Massachusetts entschieden wurden. Small Claims Courts sind zuständig für zivilrechtliche Fälle mit einem Streitwert bis \$ 1'500, das Verfahren ist mündlich und das Beweisverfahren sehr eingeschränkt, so dass die Glaubwürdigkeit der Parteien besonders wichtig ist. ZEBROWITZ/MCDONALDS nehmen an, dass unter diesen Bedingungen das Aussehen des Klägers besonders wichtig ist.¹¹³⁴ Gemäss der Feldstudie von ZEBROWITZ/MCDONALD erhöht die Attraktivität des Klägers die Wahrscheinlichkeit, dass er gewinnt, geringfügig, hat aber keinen Einfluss auf die zugesprochene Summe.¹¹³⁵ Die Attraktivität des Beklagten hat keinen signifikanten Einfluss auf den Ausgang des Verfahrens. Jedoch verlieren Beklagte mit einem erwachsenen Gesicht 92 % der auf (angeblich) vorsätzlicher Schädigung beruhenden Klagen gegen sie, aber nur 45 % der Beklagten (egal welchen Alters) mit einem Gesicht, das dem Kindchen-Schema entspricht („babyface“).¹¹³⁶ Klagen, die dem Beklagten Fahrlässigkeit vorwerfen, verlieren babygesichtige Beklagte jedoch eher. ZEBROWITZ/MCDONALD führen dies darauf zurück, dass vorsätzliche Schädigungen nicht dem Stereotyp der Naivität entsprechen, das vom Kindchen-Schema ausgelöst wird.¹¹³⁷
- 724 Zusammengefasst gesagt werden gemäss diesen beiden Feldstudien unattraktive Angeklagte zwar nicht häufiger, aber zu härteren Strafen, verurteilt, während attraktive Kläger zwar nicht mehr, aber öfter gewinnen.

III. Eigene Studie zum Attraktivitäts-Hofeffekt

A. Frage zum Attraktivitäts-Hofeffekt

- 725 Den Richterinnen und Richtern der Kantone beider Basel, Bern und Graubünden wurde folgender Sachverhalt zur Beurteilung vorgelegt (2004):

Während die Automobilistin A. sich auf einer dicht befahrenen Strasse innerorts befindet, bedient sie das Autoradio, um einen neuen Sender einzustellen.

¹¹³² JOHN E. STEWART, Defendant's Attractiveness as a Factor in the Outcome of Criminal Trials: An Observational Study, *Journal of Applied Social Psychology* 1980, 348-361, 351.

¹¹³³ STEWART, FN 1132, 359.

¹¹³⁴ LESLIE A. ZEBROWITZ/SUSAN M. McDONALD, The Impact of Litigants' Baby-Facedness and Attractiveness on Adjudications in Small Claims Courts, *Law and Human Behavior* 1991, 603-623, 606.

¹¹³⁵ ZEBROWITZ/MCDONALD, FN 1134, 620.

¹¹³⁶ ZEBROWITZ/MCDONALD, FN 1134, 614; für eine Definition des Kindchen-Schemas siehe ZEBROWITZ/MCDONALD, 605 (grosse Augen, dünne Augenbrauen, hohe Stirn, kleines Kinn, runde Gesichtsform).

¹¹³⁷ ZEBROWITZ/MCDONALD, FN 1134, 619.

Besonderer Teil

Sie übersieht deshalb, dass das Auto vor ihr angehalten hat, um dem Fussgänger F. das Überqueren der Strasse auf dem Fussgängerstreifen zu erlauben.

Der von der Automobilistin A. gelenkte Wagen prallt in das Heck des vor ihr stehenden Autos. Dieses wird durch die Wucht des Aufpralls nach vorne geschoben und verletzt den Fussgänger F. leicht. Die beiden Fahrzeuglenker bleiben unverletzt.

- 726 In der „attraktiven“ Gruppe war dem Sachverhalt ein schwarz-weisses Passfoto einer attraktiven Frau beigelegt, in der „unattraktiven“ Gruppe dasjenige einer unattraktiven Frau.¹¹³⁸ Die beiden Frauen waren auf der Website hotornot.com, wo sie ihre Bilder zur Bewertung durch die Besucher der Website veröffentlicht hatten, auf einer Skala von 1-10 mit 2,7 resp. 9,6 bewertet worden (beide mit mehr als 200 Bewertungen).¹¹³⁹ Sie sind zudem ungefähr gleich alt und zeigen den gleichen Gesichtsausdruck. Keines der beiden Gesichter entspricht dem Kindchen-Schema. Frauen wurden gewählt, weil der Schönheits-Hofeffekt bei der Beurteilung durch eine Person des anderen Geschlechts stärker ist und eine Mehrheit der Richter in der Schweiz männlich sind.
- 727 Die Richterinnen und Richter wurden gebeten, das Verschulden der Automobilistin auf einer Skala von 1-7 (1 = geringes Verschulden, 7 = schweres Verschulden) zu beurteilen sowie eine angemessene Strafe auszufällen, wobei folgende Optionen zur Auswahl standen:
- Busse in der Höhe von Fr. _____
 - Bedingte Freiheitsstrafe von ____ Tagen.
 - Busse in der Höhe von Fr. _____ und bedingte Freiheitsstrafe von ____ Tagen.
- 728 Weiter wurde darauf aufmerksam gemacht, dass über administrative Massnahmen, wie insbesondere Führerscheinentzug, nicht zu entscheiden ist (was der tatsächlichen Rechtslage entspricht).
- 729 Die Hypothese ist, dass das Verschulden der unattraktiven Automobilistin als grösser beurteilt und sie härter bestraft würde.

B. Resultate

- 730 Die Manipulation der Attraktivität zeitigte keinen signifikanten Effekt auf die Beurteilung des Verschuldens. Das Verschulden der attraktiven Autofahrerin wurde im Schnitt mit 4,28 (N = 107), dasjenige der unattraktiven mit 4,11 Punkten (N = 116) beurteilt; der geringe Unterschied ist statistisch nicht signifikant.¹¹⁴⁰
- 731 Die Art der ausgesprochenen Strafe ist in der Tabelle 19 dargestellt.

¹¹³⁸ Die Bilder sind in Anhang C abgedruckt.

¹¹³⁹ Für ein weiteres Beispiel der Verwendung von hotornot.com in der Attraktivitätsforschung siehe AMY PERFORNS, What's in a Name? – The Effect of Sound Symbolism on Perception of Facial Attractiveness, 2004, erhältlich unter www.mit.edu/~perfors/hotornot.pdf (besucht am 5. Februar 2005).

¹¹⁴⁰ Mann-Whitney U = 5677.5, Wilcoxon W = 12'463.5, p = 0,265.

		Art der Bestrafung				Total
		keine Bestrafung	nur Gefängnis	nur Busse	Busse und Gefängnis	
Gruppe	attraktiv	8	4	78	22	112
	unattraktiv	4	12	76	29	121
Total		12	16	154	51	233

Tabelle 19: Art der Bestrafung nach Gruppe

- 732 Die Unterschiede zwischen den Gruppen sind auch hier gering und statistisch nicht signifikant. Jedoch wurden von den 112 attraktiven Fahrerinnen nur 26 (23 %) zu Freiheitsstrafen (in Kombination mit einer Geldstrafe oder als alleinige Strafe) verurteilt, während 41 der 121 unattraktiven Fahrerinnen (34 %) zu bedingten Freiheitsstrafen verurteilt wurden. Dieser Effekt ist ganz knapp nicht signifikant auf dem 5 % Level.¹¹⁴¹ Wenn sie zu einer Freiheitsstrafe verurteilt wurden, war die Strafe bei den attraktiven Fahrerinnen leicht höher (26,7 Tage) als bei den unattraktiven Fahrerinnen (23,3 Tage); dieser Unterschied ist aber nicht signifikant.¹¹⁴²
- 733 Auf die Höhe der Busse hatte die Attraktivität ebenfalls keinen Einfluss (Tabelle 20). Die Busse betrug im Schnitt bei der unattraktiven Fahrerinnen Fr. 933, bei der attraktiven Fahrerinnen Fr. 939, resp. Fr. 1'000 und Fr. 600, wenn sie mit einer bedingten Freiheitsstrafe kombiniert wurde. 18 Richterinnen oder Richter meinten (zu Recht), dass sich die Höhe der Busse nach dem Einkommen bemessen würde, wobei 17 dieser Richter in der Gruppe „attraktiv“ waren. Sechs Mal wurde eine Busse in der Höhe eines Drittels eines Monatslohns als angemessenen erachtet.

		Art der Bestrafung									
		nur Busse					Busse und Gefängnis				
		Max.	Median	Min.	Durchschnitt	Standard-abw.	Max.	Median	Min.	Durchschnitt	Standard-abw.
Gruppe	attraktiv	5'000	750	200	939	787	5'000	1'000	300	1'443	1'572
	unattraktiv	5'000	900	200	933	812	2'000	600	200	772	446

Tabelle 20: Bussenhöhe in Franken

C. Diskussion

- 734 Die Manipulation der Attraktivität zeigt nicht den vermuteten Einfluss auf das Verschulden und Strafmass. Der einzige Unterschied, der statistischer Signifikanz nahe kommt, ist derjenige zwischen dem Anteil Fahrerinnen, die zu Freiheitsstrafen verurteilt wurden, der bei den unattraktiven Fahrerinnen höher ist (34 % gegenüber 33 %).
- 735 Im Nachhinein stechen zwei Schwächen des Versuchsplans ins Auge: erstens wurde es unterlassen, das Einkommen der Angeklagten anzugeben. Das Einkommen ist aber, wie einige Richter völlig zu Recht bemerkten, ein entscheidendes Kriterium für die Bussen-

¹¹⁴¹ $\chi^2 = 3,232$; $p = 0,072$.

¹¹⁴² $t = 0,553$, $p = 0,582$.

höhe (Art. 48 Ziff. 2 Abs. 2 StGB). Zweitens gab es im (sehr knappen) Sachverhalt zu wenig Anhaltspunkte dafür, dass der Unfall auf externe Faktoren zurückzuführen sein könnte. Wenn das attributionstheoretische Modell zutrifft, sollte der positive Eindruck von einer Person dazu führen, dass der Urteilende nach externen Faktoren sucht, die das (negative) Verhalten der Person erklären können. Im geschilderten Sachverhalt gab es schlicht keine Anhaltspunkte für solche externen Faktoren. Möglicherweise wäre das Resultat anders ausgefallen, wenn es Anhaltspunkte für externe Faktoren gegeben hätte – beispielsweise „Nachdem ein Fussgänger sehr überraschend auf den Fussgängerstreifen trat, konnte das vorausfahrende Auto gerade noch halten, wurde aber von der Automobilistin A. gerammt, die nicht mehr rechtzeitig auf die Vollbremsung des Vordermannes reagieren konnte“. Dieser Sachverhalt lässt offen, ob der Unfall eher auf die Unaufmerksamkeit der Angeklagten zurückzuführen ist oder auf das überraschende Verhalten des Fussgängers.

- 736 Denkbar ist, dass kulturelle Unterschiede eine Rolle spielen. Darauf deutet hin, dass SCHMITT die Resultate von LANDY und SIGALL mit deutschen Versuchspersonen nicht replizieren konnte und PIEHL bei fahrlässiger Tötung mit deutschen Versuchspersonen ebenfalls keinen Attraktivitäts-Effekt fand; dies im Gegensatz zu IZZETT/LEGINSKI mit amerikanischen Versuchspersonen.¹¹⁴³ Wenn Unterhaltungsmedien einen Einfluss auf den Schönheits-Hofeffekt haben,¹¹⁴⁴ ist es zumindest plausibel, dass in einem Land mit noch grösserer Fixierung auf die Schönen und Erfolgreichen wie den USA auch der Schönheits-Hofeffekt stärker ist. Schliesslich muss man konstatieren, dass zwar die Mehrheit der Experimente einen Einfluss der Attraktivität des Angeklagten auf das Strafmass feststellten, aber eben nicht alle. Der *attractiveness-leniency* Effekt scheint von subtilen Manipulationen des Sachverhalts und der Versuchsumstände abzuhängen, die noch nicht vollständig verstanden werden.

¹¹⁴³ SCHMITT, FN 1081, 479; PIEHL, FN 1117, 554; IZZETT/LEGINSKI, FN 1117, 275.

¹¹⁴⁴ FEINGOLD, FN 1093, 333.

§ 16 Kontrast- und Kompromisseffekt

- 737 Ein Prinzip normativer Entscheidungstheorien ist, dass eine Wahl nicht vom Vorhandensein weiterer Alternativen, die nicht gewählt werden, abhängig sein sollte (*independence of irrelevant alternatives*).¹¹⁴⁵ Dies folgt unmittelbar aus der auf S. 31 dargestellten Annahme der rationalen Entscheidungstheorie, dass jeder Mensch alle Wahlmöglichkeiten (Optionen) nach ihrer Präferenz geordnet hat und – wenn er die Wahl zwischen mehreren Optionen hat – diejenige Option wählt, die in seiner „Präferenz-Rangliste“ zuoberst steht.
- 738 Ein weiteres Prinzip der meisten normativen Entscheidungstheorien ist, dass sich der „Marktanteil“ einer Option nie erhöhen kann, wenn den zur Auswahl stehenden Optionen eine weitere Option hinzugefügt wird.¹¹⁴⁶ Das so genannte Prinzip der Regularität (*regularity*) ist ebenfalls ein Ausfluss der individuellen Nutzenmaximierung. Seine Verletzung lässt sich in aggregierten Daten beobachten und bedeutet, dass einzelne Individuen sich nicht gemäss der rationalen Entscheidungstheorie verhalten.¹¹⁴⁷ Auch Regularität ist ein Prinzip, das unmittelbar einleuchtet: wenn in einer Wettbewerbssituation zwei Produkte A und B um die Gunst der Konsumenten buhlen und ein weiteres Produkt C auf den Markt gebracht wird, ist anzunehmen, dass die Produkte A und B einen Teil ihres bisherigen Marktanteils an den neuen Konkurrenten C verlieren werden. Dass z.B. das Produkt B durch die Einführung von C an Marktanteil gewinnt, wäre sehr überraschend.¹¹⁴⁸ Dieses überraschende Resultat kann aber empirisch nachgewiesen werden und hängt von den Eigenschaften der neuen Option C ab.

I. Kontrast- und Anziehungseffekt

- 739 Kontrasteffekte in der Wahrnehmung sind allgegenwärtig und verschwinden selbst dann nicht vollständig, wenn sie dem Urteilenden bewusst gemacht werden.¹¹⁴⁹ Die meisten Menschen werden beispielsweise den Kreis links in der Abbildung 17, der von grossen Kreisen kontrastiert wird, als kleiner wahrnehmen als den Kreis rechts, der von kleineren Kreisen umgeben ist (beide Kreise sind natürlich gleich gross).

¹¹⁴⁵ AMOS TVERSKY/ITAMAR SIMONSON, Context-Dependent Preferences, *Management Science* 1993, 117-185; gekürzter Nachdruck in: KAHNEMAN/TVERSKY (Hrsg.), 518-527, 518. Die Seitenzahlen werden gemäss Nachdruck zitiert.

¹¹⁴⁶ JOEL HUBER/JOHN W. PAYNE/CHRISTOPHER PUTO, Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis, *Journal of Consumer Research* 1982, 90-98, 90 f.

¹¹⁴⁷ TVERSKY/SIMONSON, FN 1145, 519.

¹¹⁴⁸ Möglich ist natürlich, dass B bei der Einführung von C weniger Marktanteil verliert als A. Während die Annahme der *Proportionalität* davon ausgeht, dass A und B im Verhältnis ihrer bisherigen Marktanteile an C verlieren werden, postuliert die Ähnlichkeits-Hypothese (*similarity hypothesis*), dass dasjenige der alten Produkte mehr Marktanteil verliert, dass dem neuen Produkt ähnlicher (damit eher substituierbar) ist. Siehe dazu HUBER/PAYNE/PUTO, FN 1146, 91.

¹¹⁴⁹ STEWART SHAPIRO/MARK T. SPENCE, Mind Over Matter? The Inability to Counteract Contrast Effects Despite Conscious Effort, *Psychology & Marketing* 2005, 225-245.

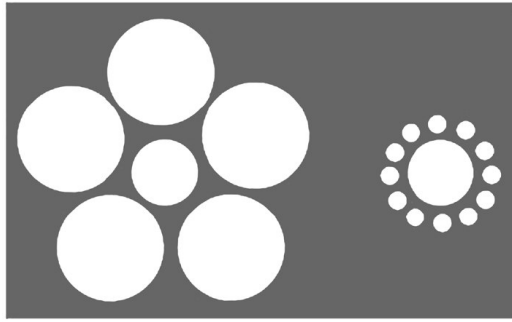


Abbildung 17: Ebbinghaus-Illusion

A. Kontrasteffekt bei asymmetrisch dominierten Alternativen

- 740 JOEL HUBER, JOHN W. PAYNE und CHRISTOPHER PUTO haben 1982 erstmals empirisch nachgewiesen, dass Kontrasteffekte nicht nur die Wahrnehmung, sondern auch die Bewertung und die Wahl von Entscheidungsalternativen beeinflussen.
- 741 Eine asymmetrisch dominierte Alternative ist eine Alternative, die von mindestens einer, aber nicht von allen zur Auswahl stehenden Alternativen dominiert wird.¹¹⁵⁰ Asymmetrische Dominanz ist in Abbildung 18 schematisch dargestellt für drei Alternativen, die sich in zwei Dimensionen (Attributen oder Eigenschaften) unterscheiden. Dies kann z.B. Preis und Qualität des Produktes sein. Das Produkt „Mitbewerber“ ist schlechter als das Produkt „Ziel“ in der Dimension 1 (z.B. Preis, also teurer), aber besser in der Dimension 2 (z.B. Qualität), während das Ziel besser als der Mitbewerber bezüglich der Dimension 1, aber schlechter bezüglich Dimension 2 ist. Keine dieser Alternativen dominiert die andere; jede hat Eigenschaften, bezüglich derer sie der anderen Alternative überlegen ist. Plastisch ausgedrückt hat der Konsument die Wahl zwischen einem teureren, aber hochwertigen Produkt und einer billigeren, aber minderwertigeren Alternative.

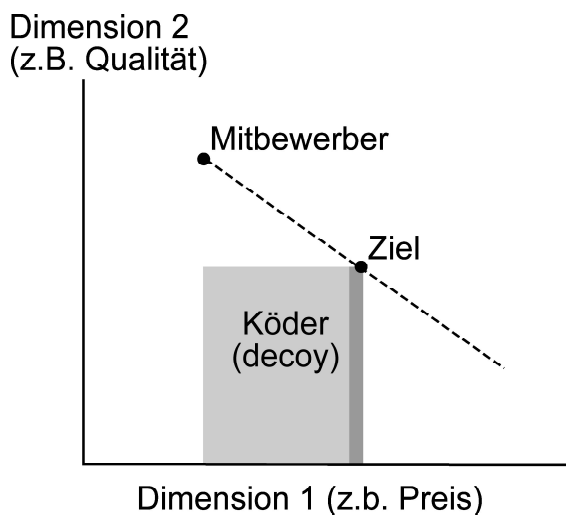


Abbildung 18: Schematische Darstellung einer asymmetrisch dominierten Alternative

¹¹⁵⁰ HUBER/PAYNE/PUTO, FN 1146, 90.

- 742 Ein weiteres Produkt, das im grauen Rechteck in der Abbildung 18 angesiedelt wird, wird jedoch vom Ziel dominiert. Das Ziel ist besser in beiden Dimensionen, sowohl bezüglich Qualität wie Preis. Hingegen wird das Produkt im grauen Bereich vom Mitbewerber nicht dominiert; der Mitbewerber ist besser bezüglich Qualität. Alle Alternativen, die sich im grauen Bereich der Abbildung 18 befinden, werden daher asymmetrisch, d.h. nur von einer der beiden anderen Alternativen, dominiert. Kein vernünftiger Mensch wird sie wählen, da das Ziel immer die bessere Wahl ist.
- 743 HUBER et al. können empirisch nachweisen, dass der Marktanteil des Ziels mit der Einführung einer asymmetrisch dominierten Alternative, die von ihnen als Köder (*decoy*) bezeichnet wird, steigt.¹¹⁵¹ Haben die Versuchspersonen beispielsweise die Wahl zwischen zwei Biersorten A (Preis \$ 2,60, Qualität „70“, Ziel) und B (Preis \$ 1,80, Qualität „50“, Mitbewerber), so wählen 43 % die Biersorte A. Steht auch noch eine dritte Sorte C (Preis \$ 3,00, Qualität „70“, Köder) zur Auswahl, so wählen 63 % die Biersorte A.¹¹⁵² Der Marktanteil des Ziels konnte trotz Einführung eines weiteren Produkts gesteigert werden, was die Regularität verletzt. Der Kontrasteffekt ist dabei gemäss HUBER et al. am stärksten, wenn der Köder im Bereich des dunkelgrauen Streifens in Abbildung 18 angesiedelt ist; d.h. wenn er sich in nur einer Dimension, und zwar derjenigen, in der das Ziel dem Mitbewerber unterlegen ist, vom Ziel unterscheidet.¹¹⁵³ Als Kontrasteffekt wird das Phänomen bezeichnet, weil das Ziel im Kontrast zum dominierten Köder attraktiver erscheint.¹¹⁵⁴
- 744 Der Kontrasteffekt wurde durch verschiedene Forscher repliziert und gilt als robust.¹¹⁵⁵ DOYLE et al. konnten ihn auch ausserhalb des Labors nachweisen. In ihrem Experiment standen in einem Lebensmittelladen in der ersten Woche zwei Sorten von *Baked Beans* zur Auswahl: die HEINZ Dose à 420 g für 29 Pence und die SPAR Dose à 420 g für 21 Pence. Konsumentenbefragungen hatten ergeben, dass die Qualität von HEINZ Baked Beans als besser gilt. In der Woche 1 betrug der Marktanteil von SPAR 19 %, d.h. die meisten Konsumenten zogen die höherpreisige, qualitativ bessere Alternative vor. In der Woche 2 wurde ein Köder gelegt: eine Dose SPAR Baked Beans à 220 g für 21 Pence. Der Köder ist mit anderen Worten gleich teuer wie die 420 g SPAR Dose, hat aber nur (ungefähr) die Hälfte ihres Inhalts; er wird vom Ziel dominiert. Wenig überraschend wurde denn auch keine einzige Dose SPAR à 220 g verkauft. Aber: der Marktanteil der 420 g SPAR Dose betrug in der zweiten Woche 33 %, eine Steigerung um 14 Prozentpunkte – dies entspricht einer Absatzsteigerung von 73 %!¹¹⁵⁶
- 745 Die Verletzung der Regularität bei der Einführung asymmetrisch dominierter Alternativen konnte auch bei Honigbienen (*Apis mellifera*) und einer kanadischen Vogelart (*Perisoreus canadensis*) beobachtet werden. Den Singvögeln wurde zuerst Futter in zwei Röhren ange-

¹¹⁵¹ HUBER/PAYNE/PUTO, FN 1146, 92.

¹¹⁵² HUBER/PAYNE/PUTO, FN 1146, 95, 97.

¹¹⁵³ HUBER/PAYNE/PUTO, FN 1146, 94.

¹¹⁵⁴ ITAMAR SIMONSON/AMOS TVERSKY, Choice in Context: Tradeoff Contrast and Extremeness Aversion, Journal of Marketing Research 1992, 281-295, 282.

¹¹⁵⁵ JOHN R. DOYLE/DAVID J. O'CONNOR/GARETH M. REYNOLDS/PAUL A. BOTTOMLEY, The Robustness of the Asymmetrically Dominated Effect: Buying Frames, Phantom Alternatives, and In-Store Purchases, Psychology & Marketing 1999, 225-243, 228 m.H.

¹¹⁵⁶ DOYLE/O'CONNOR/REYNOLDS/BOTTOMLEY, FN 1155, 240.

boten; in der einen Röhre war das Futter 28 cm vom Eingang der Röhre entfernt; in der anderen Röhre die doppelte Menge Futter in der doppelten Distanz (Ziel). Anschliessend wurde eine dritte Röhre im Abstand von 84 cm vom Eingang mit der gleichen Menge Futter bestückt, die in der Zielröhre war (Köder). Der Köder wird vom Ziel asymmetrisch dominiert, weil der Weg zum Futter länger ist, ohne dass mehr Futter angeboten würde. Der Köder wurde denn auch nur in 3,8 % aller Versuche gewählt. Aber das Ziel wurde nach Einführung des Köders zulasten des Mitbewerbers signifikant häufiger gewählt; die Reihenfolge der Häufigkeit der Wahl von Mitbewerber und Ziel kehrte sich um, wenn der Köder zur Auswahl stand.¹¹⁵⁷

- 746 Während die durch den Kontrasteffekt bewirkte Verschiebung des Marktanteils zugunsten des Ziels relativ gering ist, wenn den gleichen Personen zuerst die Auswahl zwischen Mitbewerber und Ziel und später (nach ein bis zwei Wochen) die Wahl zwischen Mitbewerber, Ziel und Köder angeboten wird,¹¹⁵⁸ ist sie erheblich, wenn einer Gruppe die Wahl zwischen Mitbewerber und Ziel und einer anderen Gruppe zwischen Mitbewerber, Ziel und Köder angeboten wird.¹¹⁵⁹ HUBER et al., DOYLE et al. und SIMONSON/TVERSKY berichten übereinstimmend von einer Verschiebung zugunsten des Ziels von 10-15 Prozentpunkten, was – je nach ursprünglichem Marktanteil des Ziels – einer Steigerung des Marktanteils um 30-50 % entspricht.¹¹⁶⁰ Der Kontrasteffekt ist nicht nur signifikant, er ist auch erheblich.

B. Anziehungseffekt bei klar minderwertigen Alternativen

- 747 JOEL HUBER und CHRISTOPHER PUTO haben den bei der Einführung asymmetrisch dominierter Alternativen beobachteten Kontrasteffekt 1983 auf klar minderwertige Alternativen erweitert.¹¹⁶¹ Eine klar minderwertige Alternative wird von keiner anderen Alternative dominiert, erscheint aber gegenüber dem Ziel als deutlich weniger begehrenswert.¹¹⁶² Die Erweiterung des Kontrasteffekts bei asymmetrisch dominierten Alternativen auf den Anziehungseffekt bei klar minderwertigen Alternativen ist praktisch wichtig, weil asymmetrisch dominierte Alternativen im wirklichen Leben selten, klar minderwertige Alternativen aber häufig vorkommen. Wenn sich auch die Einführung einer klar minderwertigen Alternative zugunsten des Ziels auswirkt, ist die praktische Bedeutung der unter Laborbedingungen beobachteten Phänomene erheblich grösser.

¹¹⁵⁷ SHARONI SHAFIR/TOM A. WAITE/BRIAN H. SMITH, Context-Dependent Violations of Rational Choice in Honeybees (*Apis Mellifera*) And Gray Jaws (*Perisoreus Canadensis*), *Behavioral Ecology and Sociobiology* 2002, 180-187, 183 f., 185.

¹¹⁵⁸ Sogenanntes *within-subject design*; HUBER/PAYNE/PUTO, FN 1146, 94, massen für das *within-subject design* eine Erhöhung des Marktanteils des Ziels von 53 % auf 56 %; ein zwar geringer, aber statistisch signifikanter Effekt.

¹¹⁵⁹ Sogenanntes *between-subject design*.

¹¹⁶⁰ HUBER/PAYNE/PUTO, FN 1146, 95; DOYLE/O'CONNOR/REYNOLDS/BOTTOMLEY, FN 1155, 233; SIMONSON/AMOS TVERSKY, FN 1154, 287.

¹¹⁶¹ JOEL HUBER/CHRISTOPHER PUTO, Market Boundaries and Product Choice: Illustrating Attraction and Substitution Effects, *Journal of Consumer Research* 1983, 31-44.

¹¹⁶² HUBER/PUTO, FN 1161, 33.

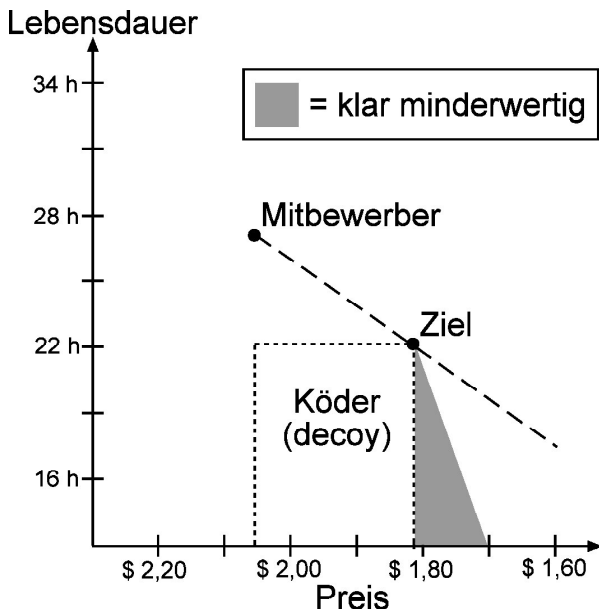


Abbildung 19: Schematische Darstellung einer klar minderwertigen Alternative

- 748 Abbildung 19 zeigt schematisch, in welchem Bereich eine klar minderwertige Alternative angesiedelt ist. Setzt man den Trade-off zwischen Dimension 1 und Dimension 2 linear fort (gestrichelte Linie in Abbildung 19), so liegt die klar minderwertige Alternative deutlich unter dieser Trendlinie, im grauen Bereich der Abbildung 19. Beispielsweise kann der Mitbewerber eine Batterie mit einer Lebensdauer von knapp 28 Stunden sein, die \$ 2,05 kostet, und das Ziel eine Batterie mit einer Lebensdauer von 22 Stunden, die nur \$ 1,80 kostet. Mit anderen Worten kostet jede zusätzliche Stunde Lebensdauer rund \$ 0,042 (vier Cent). Eine minderwertige Alternative im grauen Bereich würde beispielsweise bei einer Lebensdauer von 16 Stunden \$ 1,75 kosten, was erheblich mehr ist als bei einer linearen Beziehung von Lebensdauer und Preis (gemäss der eine Batterie mit einer Lebensdauer von 16 Stunden \$ 1,55 kosten würde). Das Konzept der minderwertigen Alternative wird von HUBER und PUTO zwar nicht genau definiert, für die meisten praktischen Anwendungen ist aber für jedermann klar, was eine minderwertige Alternative ist.¹¹⁶³
- 749 Die Einführung einer klar minderwertigen Alternative wirkt sich gemäss HUBER und PUTO ähnlich wie die Einführung einer asymmetrisch dominierten Alternative auf den Marktanteil des Ziels aus. Die minderwertige Alternative zieht die getroffene Wahl in Richtung des Ziels, das der minderwertigen Alternative am ähnlichsten erscheint; HUBER und PUTO sprechen deshalb von einem Anziehungseffekt (*attraction effect*).¹¹⁶⁴

	Alternativen (Batterien)			
	Köder 1	Alternative 1	Alternative 2	Köder 2
Lebensdauer →	14 Stunden	22 Stunden	28 Stunden	32 Stunden
Preis →	\$ 1,50	\$ 1,80	\$ 2,10	\$ 2,70
Gruppe 1	Nicht erhältl.	2	10	3
Gruppe 2	1	10	6	Nicht erhältl.

Tabelle 21: Resultate der Studie 2 von HUBER und PUTO (1983)

¹¹⁶³ DOYLE/O'CONNOR/REYNOLDS/BOTTOMLEY, FN 1155, 228.

¹¹⁶⁴ HUBER/PUTO, FN 1161, 32.

- 750 Die Tabelle 21 zeigt beispielhaft, was beim Markteintritt des Köders passiert. Die Wahl zwischen den Alternativen 1 und 2 verschiebt sich jeweils zugunsten derjenigen Alternative, der eine minderwertige Alternative gegenübergestellt wird.¹¹⁶⁵ Je weniger minderwertig der Köder ist, d.h. je eher Köder und Ziel substituierbar sind, desto mehr des ursprünglichen Marktanteils des Ziels übernimmt der Köder.¹¹⁶⁶ Der Anziehungseffekt lässt sich daher am besten beobachten, wenn der Köder klar minderwertig ist und kaum gewählt wird. Wird der Köder regelmässig gewählt, vermischt sich die Wirkung des Anziehungseffekts (wirkt zu Gunsten des Ziels) mit dem Substitutionseffekt (wirkt zu Lasten des Ziels), so dass sich der beobachtete aggregierte Effekt kaum mehr von einer proportionalen Verschiebung der Marktanteile zugunsten der neu eingeführten Alternative, wie sie von klassischen Modellen vorhergesagt wird, unterscheiden lässt.¹¹⁶⁷

II. Kompromisseffekt (auch *extremeness aversion*)

- 751 Der Kompromisseffekt wurde erstmals 1989 von ITAMAR SIMONSON beschrieben.¹¹⁶⁸ Er besagt, dass bei einer Wahl zwischen drei Alternativen in der Regel die mittlere Alternative bevorzugt wird. SIMONSON erklärt dies damit, dass sich der Entscheidende oft nicht sicher ist, welches der beiden Attribute des Produktes für ihn wichtiger ist. Die mittlere Option stellt einen Kompromiss dar, der beide Attribute kombiniert.¹¹⁶⁹ Anders als eine klar minderwertige Alternative ist die extreme Alternative nicht minderwertig, sondern liegt auf der gedachten Trendlinie, die den Trade-off der beiden Attribute darstellt, aber eben weiter entfernt von der Mitte (siehe Abbildung 20).

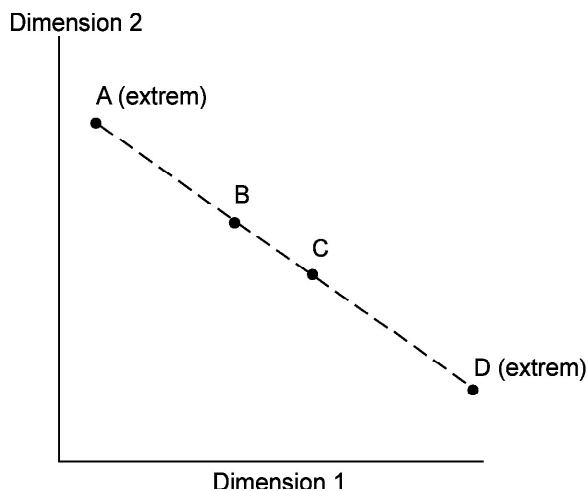


Abbildung 20: Schematische Darstellung von zwei Kern- und zwei extremen Alternativen

¹¹⁶⁵ Tabelle 21 wurde HUBER/PUTO, FN 1161, 34, entnommen und leicht angepasst.

¹¹⁶⁶ HUBER/PUTO, FN 1161, 36.

¹¹⁶⁷ HUBER/PUTO, FN 1161, 38.

¹¹⁶⁸ ITAMAR SIMONSON, Choice Based on Reason: The Case of Attraction and Compromise Effects, Journal of Consumer Research 1989, 158-174.

¹¹⁶⁹ SIMONSON, FN 1168, 161.

- 752 Eine Spiegelreflexkamera Minolta X-370 für \$ 169,99 wird z.B. von 50 % der Versuchspersonen gewählt, wenn daneben eine Minolta Maxxum 3000i für \$ 239,99 zur Auswahl steht. Ist zusätzlich noch die hochpreisige Minolta Maxxum 7000i für \$ 469,99 im Angebot, wählen 57 % der Versuchspersonen die mittlere Maxxum 3000i, 21 % die teure Maxxum 7000i und nur noch 22 % die günstige X-370.¹¹⁷⁰ Wären die Präferenzen der Versuchspersonen konstant, hätte der Marktanteil der teuren Kamera zu Lasten des mittleren Preissegments gehen müssen, weil diejenigen, die für mehr Leistung gerne etwas mehr bezahlen, von der 3000i zur 7000i wechseln (und nicht von der X-370 zur 7000i).¹¹⁷¹ Offenbar werden aber die Präferenzen durch die zur Verfügung stehenden Wahlmöglichkeiten beeinflusst. Eine weitere Demonstration des Kompromisseffekts ist in Tabelle 22 dargestellt, wo ein linearer Zusammenhang zwischen den Attributen „Grillfläche“ und „Gewicht“ eines mobilen Campinggrills besteht. Drei verschiedenen Gruppen von Versuchspersonen wurden jeweils drei der fünf Modelle zur Auswahl angeboten; die mittlere Option (in der Tabelle 22 fett) wurde jedes Mal am häufigsten gewählt.¹¹⁷²

Attribute		Marktanteil		
Kochfläche (in Quadrat-Inch)	Gewicht (in Pfund)	Gruppe 1	Gruppe 2	Gruppe 3
160	4	32 %	–	–
220	7	40 %	27 %	–
280	10	29 %	44 %	26 %
340	13	–	29 %	47 %
400	16	–	–	26 %

Tabelle 22: Beispiel für Kompromisseffekt (aus SIMONSON/TVERSKY, 1992)

- 753 Nach SIMONSON ist der Kompromisseffekt grösser, wenn der Wählende seine Wahl später gegenüber Drittpersonen rechtfertigen muss. Da der Wähler die Präferenzen der Dritten noch weniger als seine eigenen kennen kann, ist es riskant, die Wahl damit zu rechtfertigen, dass die eine Dimension viel wichtiger sei als die andere („Schatz, natürlich wiegt der Campinggrill fast zehn Kilo, aber dafür haben auch alle Steaks gleichzeitig Platz“). Die Wahl des Kompromisses erscheint als die sichere Wahl, da der maximal mögliche Fehler am geringsten ist. Die Wahl des Kompromisses lässt sich auch damit rechtfertigen, dass er beide Dimensionen optimal kombiniert.¹¹⁷³ SIMONSON hat daher der Hälfte seiner Versuchspersonen gesagt, dass ihre Wahl völlig anonym sei. Der anderen Hälfte wurde gesagt, dass alle Antworten in einer Broschüre neben ihrem Namen publiziert und diese Broschüre in der nächsten Unterrichtsstunde verteilt würde, wobei es möglich sei, dass sie ihre Wahl vor der Klasse rechtfertigen müssten. Diese Gruppe musste ihren Namen auf den Fragebogen setzen und jede Seite paraphieren.¹¹⁷⁴ Der Kompromisseffekt war in der Gruppe, die damit rechnen musste, die Wahl vor Dritten rechtfertigen zu müssen, signifikant und erheblich (rund 10-20 %) stärker als in der anonymen Gruppe.¹¹⁷⁵ Da Gerichte

¹¹⁷⁰ SIMONSON/TVERSKY, FN 1154, 290.

¹¹⁷¹ Annahme der Rangordnung der Präferenzen.

¹¹⁷² SIMONSON/TVERSKY, FN 1154, 290.

¹¹⁷³ SIMONSON, FN 1168, 162.

¹¹⁷⁴ SIMONSON, FN 1168, 163.

¹¹⁷⁵ SIMONSON, FN 1168, 165.

ihre Urteile rechtfertigen müssen, legen diese Resultate nahe, dass der Kompromisseffekt das Urteil von Richtern besonders stark beeinflusst.

- 754 TVERSKY erklärt den Kompromisseffekt mit der Abneigung gegen Verluste (*loss aversion*). Die Abneigung gegen Verluste, betrachtet vom Status quo aus, wird dabei ausgedehnt auf den Vergleich der Vor- und Nachteile der Alternativen. Wenn Nachteile stärker empfunden werden als Vorteile gleicher absoluter Höhe, so spricht dies für die Wahl der mittleren Alternative aus einer Gruppe, da die Nachteile dieser Alternative gegenüber den anderen Alternativen nur gering sind. TVERSKY/SIMONSON sprechen daher statt von Kompromisseffekt auch von *extremeness aversion*.¹¹⁷⁶
- 755 Der Kompromisseffekt wurde mit so unterschiedlichen Gütern wie Wohnungen, Vermögensanlage-Portfolios und Mundspülungen nachgewiesen und gilt als sehr robust.¹¹⁷⁷ Eine Alternative gewinnt nach übereinstimmenden Berichten von SIMONSON und KIVETZ et al. rund 15-20 % Prozentpunkte Marktanteil hinzu, wenn sie zur Kompromissoption wird.¹¹⁷⁸ KIVETZ et al. demonstrieren einen – naturgemäss in absoluten Werten etwas geringeren – Kompromisseffekt auch in komplexeren Wahlsituationen mit fünf Optionen, die sich in vier Dimensionen unterscheiden.¹¹⁷⁹

III. Beeinflussung rechtlicher Entscheidungen durch die Anzahl in Betracht gezogener Alternativen

A. Klassische Studie von Neil Vidmar

- 756 NEIL VIDMAR untersuchte bereits 1972, welchen Einfluss eine Veränderung der zur Auswahl stehenden Alternativen (des *choice set*) auf juristische Entscheidungen hat.¹¹⁸⁰ Die Studie von VIDMAR ist von einer wahren Begebenheit inspiriert. 1969 wurde ein weisser Polizist von einem Geschworenengericht in Michigan des Mordes an einem schwarzen Mann während der Rassenunruhen in Detroit freigesprochen. Der Richter instruierte die Geschworenen, dass sie nur die Möglichkeit hätten, den Angeklagten entweder freizusprechen oder wegen Mordes zu verurteilen, was automatisch lebenslänglich bedeutete. In der juristischen Literatur wurde diskutiert, ob es möglicherweise zu einer Verurteilung wegen eines geringeren Delikts gekommen wäre, wenn der Richter den Geschworenen erlaubt hätte, auch auf ein geringeres Delikt (z.B. Tötung im Notwehrexzess) zu erkennen. VIDMAR untersuchte diese Spekulation empirisch.

¹¹⁷⁶ SIMONSON/TVERSKY, FN 1154, 282.

¹¹⁷⁷ RAN KIVETZ/ODED NETZER/V. SRINIVASAN, Alternative Models for Capturing the Compromise Effect, *Journal of Marketing Research* 2004, 237-257, 238.

¹¹⁷⁸ SIMONSON, FN 1168, 166; KIVETZ/NETZER/SRINIVASAN, FN 1177, 246.

¹¹⁷⁹ KIVETZ/NETZER/SRINIVASAN, FN 1177, 252.

¹¹⁸⁰ NEIL VIDMAR, Effects of Decision Alternatives on the Verdicts and Social Perception of Simulated Jurors, *Journal of Personality and Social Psychology* 1972, 211-218.

	Gruppe						
	Eins	Zwei	Drei	Vier	Fünf	Sechs	Sieben
Mord (25 Jahre bis lebensl.)	46 %	–	–	8 %	29 %	–	8 %
Vorsätzl. Tötung (5-20 Jahre)	–	84 %	–	22 %	–	46 %	63 %
Totschlag (1-5 Jahre)	–	–	92 %	–	67 %	54 %	21 %
Freispruch	54 %	16 %	8 %	0 %	4 %	0 %	8 %

Tabelle 23: Resultate von VIDMAR (1972)

- 757 VIDMAR informierte seine Versuchspersonen nicht über die Elemente der einzelnen Tatbestände, sondern nur über die Sanktion, die jeder Tatbestand zwingend zur Folge hatte (linke Spalte in der Tabelle 23). Seine Resultate zeigen, dass die Anzahl der Freisprüche sinkt, wenn die Geschworenen auf einen Tatbestand erkennen können, der eine geringere Strafe zur Folge hat. Die Resultate von VIDMAR wurden später von HAMILTON repliziert und von KOCH/DEVINE auf Gruppen, die sich vor dem Urteil beraten, erweitert.¹¹⁸¹
- 758 Die Studie von VIDMAR ist interessant, aber sie belegt keine Verletzung der Regularität. Keine der beiden ursprünglichen Alternativen (Mord und Freispruch) gewinnt durch die Hinzufügung weiterer Alternativen an „Marktanteil“. Die Resultate lassen sich zwanglos durch individuelle Nutzenmaximierung erklären. Wenn ein Geschworener beispielsweise der Ansicht ist, dass eine Freiheitsstrafe von fünf Jahren dem Verschulden des Angeklagten angemessen ist, wird er den Angeklagten freisprechen, wenn nur Freispruch und 25 Jahre Gefängnis zur Auswahl stehen. Der Freispruch liegt näher an der seiner Meinung nach optimalen Strafe; er hat mit anderen Worten die Wahl, den Angeklagten 20 Jahre zu viel oder fünf Jahre zu wenig ins Gefängnis zu schicken. Die Gewissensbisse – und somit die Disutilität – sind im ersten Fall wahrscheinlich grösser.¹¹⁸²
- 759 Die Studie von VIDMAR hat in der Schweiz im Strafrecht keine grosse Bedeutung, weil das Gericht nicht an die rechtliche Würdigung der Anklage durch die Anklagebehörde gebunden ist (*iura novit curia*).¹¹⁸³ Sie ist jedoch für das Zivilrecht, wo das Gericht grundsätzlich an das Rechtsbegehren des Klägers gebunden ist, durchaus interessant.¹¹⁸⁴ Jeder Anwalt, der ein Rechtsbegehren mit Eventualbegehren formuliert, überlegt sich, ob der Einschluss der Eventualstandpunkte die Chance, dass das Hauptbegehren gutgeheissen wird, verringert. Die Studie von VIDMAR legt nahe, dass die Wahrscheinlichkeit, das Hauptbegehren zuerkannt zu erhalten, durch die Nennung von Eventualbegehren tatsächlich sinkt, aber dass dafür die Chance, dass die Klage zumindest teilweise gutgeheissen wird, steigt.

¹¹⁸¹ HAMILTON (1976), zitiert bei BERNARD GROFMAN, *Mathematical Models of Juror and Jury Decision-Making*, in: BRUCE DENNIS SALES (Hrsg.), *The Trial Process*, New York 1981, 305-352, 318 f.; CHANTAL MEES KOCH/DENNIS J. DEVINE, *Effects of Reasonable Doubt Definition and Inclusion of a Lesser Charge on Jury Verdicts*, *Law and Human Behavior* 1999, 653-674.

¹¹⁸² Weitere Erklärungen der Daten von VIDMAR durch GROFMAN, FN 1181, 318 ff.; KINLEY LARNTZ, *Reanalysis of Vidmar's Data on the Effects of Decision Alternatives on Verdicts of Simulated Jurors*, *Journal of Personality and Social Psychology* 1975, 123-140.

¹¹⁸³ SCHMID, FN 790, Rz. 149.

¹¹⁸⁴ Das Gericht kann eine Klage natürlich nur teilweise gutheissen, aber es kann auf nichts anderes als auf das Begehrte erkennen.

B. Demonstration von Kontrast- und Kompromisseffekt im Recht durch Kelman et al.

- 760 KELMAN, ROTTENSTREICH und TVERSKY waren die ersten, die den Kompromiss- und Kontrasteffekt bei spezifisch juristischen Entscheidungen demonstrieren konnten.¹¹⁸⁵
- 761 In einem ihrer Experimente wurde den Versuchspersonen geschildert, dass der schwarze Angeklagte von einem weissen Angestellten eines privaten Sicherheitsdienstes zu Unrecht des Einbruchs verdächtigt worden sei. In der Folge kam es zum Streit zwischen dem Angeklagten und dem Wachmann, der den Angeklagten rassistisch beschimpfte, worauf der Angeklagte eine Pistole zückte und den Wachmann erschoss (stark verkürzter Sachverhalt).¹¹⁸⁶ Den Versuchspersonen wurde erklärt, dass qualifizierter Mord, Mord, vorsätzliche Tötung und fahrlässige Tötung als mögliche Straftatbestände in Frage kämen.¹¹⁸⁷ Fahrlässige Tötung käme nur in Frage, wenn der Angeklagte irrtümlicherweise gemeint habe, sich verteidigen zu dürfen. Die Unterscheidung zwischen Mord und qualifiziertem Mord hänge davon ab, ob der Wachmann ein Polizist im Sinne des qualifizierten Tatbestandes sei. Auf vorsätzliche Tötung statt auf Mord könne erkannt werden, wenn das Opfer den Angeklagten zulänglich provoziert habe.¹¹⁸⁸
- 762 Der einen Gruppe der Versuchspersonen wurde anschliessend gesagt, dass der Richter entschieden haben, ein Wachmann sei von Rechts wegen kein Polizist im Sinne des qualifizierten Tatbestandes, weshalb eine Verurteilung wegen qualifizierten Mordes ausgeschlossen sei. Der anderen Gruppe wurde gesagt, der Richter hätte entschieden, dass fahrlässige Tötung aus rechtlichen Gründen ausgeschlossen sei. Beiden Versuchsgruppen wurden also zuerst alle vier möglichen Tatbestände geschildert, aber dann wurde das *choice set* so eingeschränkt, dass jeweils ein anderer Tatbestand in der Mitte war und gemäss dem Kompromisseffekt häufiger gewählt werden müsste.

<i>Tatbestände</i>	<i>Obere Gruppe</i>	<i>Untere Gruppe</i>
Qualifizierter Mord	13 %	–
Mord	57 %	38 %
Vorsätzliche Tötung	30 %	55 %
Fahrlässige Tötung	–	7 %

Tabelle 24: Resultate der Studie 2 von Kelman et al. (1996)

- 763 Die Resultate von KELMAN et al. demonstrieren einen klaren Kompromisseffekt. Die mittlere Option wird in beiden Gruppen am häufigsten gewählt. Dadurch erhöht sich der Anteil der Verurteilungen für Mord von 38 % auf 55 %, wenn qualifizierter Mord als Alternative zur Auswahl steht.¹¹⁸⁹ Der Angeklagte fährt also erheblich schlechter, wenn das Gesetz

¹¹⁸⁵ MARK KELMAN/YUVAL ROTTENSTREICH/AMOS TVERSKY, Context-Dependence in Legal Decision Making, *Journal of Legal Studies* 1996, 287-318; Nachdruck in: SUNSTEIN (Hrsg.), 61-94. Die Seitenzahlen werden gemäss dem Nachdruck zitiert.

¹¹⁸⁶ KELMAN/ROTTENSTREICH/TVERSKY, FN 1185, 64 f.

¹¹⁸⁷ Im Original: special circumstances murder, murder, voluntary manslaughter, involuntary manslaughter.

¹¹⁸⁸ „Adequately provoked“, KELMAN/ROTTENSTREICH/TVERSKY, FN 1185, 65.

¹¹⁸⁹ KELMAN/ROTTENSTREICH/TVERSKY, FN 1185, 67.

einen besonders qualifizierten Tatbestand vorsieht, selbst wenn er nicht des qualifizierten Tatbestandes schuldig gesprochen wird. Der Gesetzgeber muss sich bewusst sein, dass die Einführung eines qualifizierten Tatbestandes auch die Präferenzen für die übrigen Tatbestände ändern kann.¹¹⁹⁰

IV. Die lebenslängliche Verwahrung: eine neue Alternative

764 Das revidierte Strafgesetzbuch, das am 13. Dezember 2002 verabschiedet wurde (BBl 2002, 8240), aber noch nicht in Kraft getreten ist, sieht eine zeitlich unbegrenzte Verwahrung bei Rückfallgefahr bereits für Ersttäter, die jemanden schwer geschädigt haben oder schädigen wollten, vor (Art. 64 revidiertes StGB). Zu verwahren sind psychisch unauffällige Täter (Art. 64 lit. a revStGB) und psychisch schwer gestörte Täter, bei denen eine stationäre Behandlung keinen Erfolg verspricht (Art. 64 lit. b revStGB). Die Verwahrung wird in Anschluss an die Freiheitsstrafe vollzogen, so dass der Täter mindestens während der Dauer der Freiheitsstrafe seiner Freiheit beraubt ist. Über die bedingte Entlassung muss mindestens einmal jährlich beschlossen werden (Art. 64b Abs. 1 revStGB). Entlassungen aus der Verwahrung sind nur bedingt und gestützt auf eine unabhängige sachverständige Begutachtung und nach Anhörung einer Fachkommission möglich (Art. 64b revStGB).

765 Am 8. Februar 2004 wurde die Volksinitiative „Lebenslange Verwahrung für nicht therapierbare, extrem gefährliche Sexual- und Gewaltstraftäter“ von Volk und Ständen gutgeheissen. Absätze 1 und 2 des aufgrund der Volksinitiative neu in die Bundesverfassung aufgenommenen Artikels 123a lauten:

¹Wird ein Sexual- oder Gewaltstraftäter in den Gutachten, die für das Gerichtsurteil nötig sind, als extrem gefährlich erachtet und nicht therapierbar eingestuft, ist er wegen des hohen Rückfallrisikos bis an sein Lebensende zu verwahren. Frühzeitige Entlassung und Hafturlaub sind ausgeschlossen.

²Nur wenn durch neue, wissenschaftliche Erkenntnisse erwiesen wird, dass der Täter geheilt werden kann und somit keine Gefahr mehr für die Öffentlichkeit darstellt, können neue Gutachten erstellt werden. Sollte auf Grund dieser neuen Gutachten die Verwahrung aufgehoben werden, so muss die Haftung für einen Rückfall des Täters von der Behörde übernommen werden, die die Verwahrung aufgehoben hat.

766 Die vom Bundesrat eingesetzte Arbeitsgruppe „Verwahrung“ hat zur Umsetzung der Volksinitiative eine Anpassung des revidierten Strafgesetzbuches empfohlen.¹¹⁹¹ Art. 64 Abs. 1^{ter} des Vorentwurfs der Arbeitsgruppe („Vorentwurf“) sieht neben der „ordentlichen“ die „lebenslängliche“ Verwahrung vor, wenn die folgenden Voraussetzungen erfüllt sind:

^{1ter}Das Gericht ordnet die lebenslängliche Verwahrung an, wenn der Täter einen Mord, eine vorsätzliche Tötung, eine schwere Körperverletzung, eine Vergewaltigung, einen Raub, eine Geiselnahme, eine Brandstiftung oder ein anderes Verbrechen begangen hat, durch das er die physische, psychische oder sexuelle Integrität einer anderen Person schwer beeinträchtigt hat oder beeinträchtigen wollte, und

¹¹⁹⁰ KELMAN/ROTTENSTREICH/TVERSKY, FN 1185, 75.

¹¹⁹¹ Vorentwurf der Arbeitsgruppe „Verwahrung“ vom 15. Juli 2004 zur Änderung des Strafgesetzbuches in der Fassung vom 13. Dezember 2002 betreffend die Umsetzung von Artikel 123a BV (lebenslängliche Verwahrung) und einzelne nachträgliche Korrekturen am neuen Massnahmerecht, erhältlich unter www.ofj.admin.ch/themen/lebverwahr/vorentwurf-d.pdf (besucht am 15. Februar 2005).

Besonderer Teil

- a) beim Täter im Vergleich zu anderen Tätern solcher Delikte eine sehr hohe Wahrscheinlichkeit dafür besteht, dass er ein weiteres Verbrechen dieser Art begeht, und
- b) der Täter aufgrund besonderer Persönlichkeitsmerkmale als dauerhaft nicht therapierbar eingestuft wird, weil die Behandlung langfristig keinen ausreichenden Erfolg verspricht.

- 767 Die vorzeitige bedingte Entlassung aus der Freiheitsstrafe (die ebenfalls vor der Verwahrung vollzogen wird) ist bei der lebenslänglichen Verwahrung ausgeschlossen. Anders als die Volksinitiative sieht der Vorentwurf ein Kontrollverfahren vor (Art. 64c Vorentwurf), das aber nach Auffassung namhafter Rechtslehrer den Anforderungen von Art. 5 Abs. 4 EMRK nicht genügt.¹¹⁹² Andererseits werden die Anforderungen an die „normale“ Verwahrung gesenkt; es ist keine besonders qualifizierte Anlasstat mehr notwendig, sondern jedes Vergehen oder Verbrechen genügt, wenn der Täter besonders gefährlich ist (Art. 64 Abs. 1 Vorentwurf). Der Vorentwurf wurde von verschiedener Seite heftig kritisiert, so insbesondere von den Professoren und Professorinnen für öffentliches Recht der Universität Bern.¹¹⁹³
- 768 Ich will mich hier nicht zur EMRK-Konformität oder Verfassungsmässigkeit der vorgeschlagenen Änderungen äussern. Die Einführung der „lebenslänglichen“ Verwahrung nach Art. 64 Abs. 1ter Vorentwurf erweitert auf jeden Fall die dem Gericht zur Verfügung stehenden Wahlmöglichkeiten. Die „ordentliche“ Verwahrung nach Art. 64 Abs. 1 Vorentwurf ist nun nicht mehr die extremste, die persönliche Freiheit am stärksten beschränkende Massnahme, sondern eine mittlere Lösung (auch eine „ordentliche“ Verwahrung kann lebenslänglich sein, wenn der Täter gefährlich bleibt. Es hat sich aber eingebürgert, von „lebenslänglicher“ Verwahrung zu sprechen, wenn die Verwahrung im Sinne von Art. 123a BV gemeint ist.¹¹⁹⁴ Diese Terminologie wird hier übernommen und im Folgenden auf die Anführungszeichen verzichtet).
- 769 Jede Freiheitsstrafe und die Bewegungsfreiheit einschränkende Massnahme (Verwahrung) schränkt die persönliche Freiheit (Art. 10 Abs. 2 BV) des Betroffenen ein. Die Rechtfertigung für die Einschränkung liegt in der spezialpräventiven Wirkung der freiheitsbeschränkenden Massnahme (die Strafe wird nicht ausschliesslich, aber auch, aus spezialpräventiven Gründen angeordnet). Bei gefährlichen Tätern ist das öffentliche Interesse an einer Einschränkung ihrer Bewegungsfreiheit zur Verhinderung weiterer Straftaten grösser als das private Interesse des Einzelnen, sich ungehindert bewegen zu können. Jede freiheitsbeschränkende Massnahme stellt nach dieser Betrachtungsweise einen Trade-off zwischen persönlicher Freiheit und öffentlicher Sicherheit, oder Spezialprävention, dar (Abbildung 21).

¹¹⁹² Zur Kritik an dieser Regelung siehe KARL-LUDWIG KUNZ/ANDREA BÄCHTOLD/REGINA KIENER/GÜNTER HEINE/PIERRE TSCHANNEN/WALTER KÄLIN/JÖRG KINZIG, Universität Bern zum Vorentwurf der Arbeitsgruppe „Verwahrung“, Jusletter vom 13. Dezember 2004, Rz. 7 ff., erhältlich unter www.cx.unibe.ch/krim/Vernehmlassung%20Uni.pdf, besucht am 15. Februar 2005.

¹¹⁹³ A. a. O.

¹¹⁹⁴ Z.B. bei KUNZ et al., FN 1192.

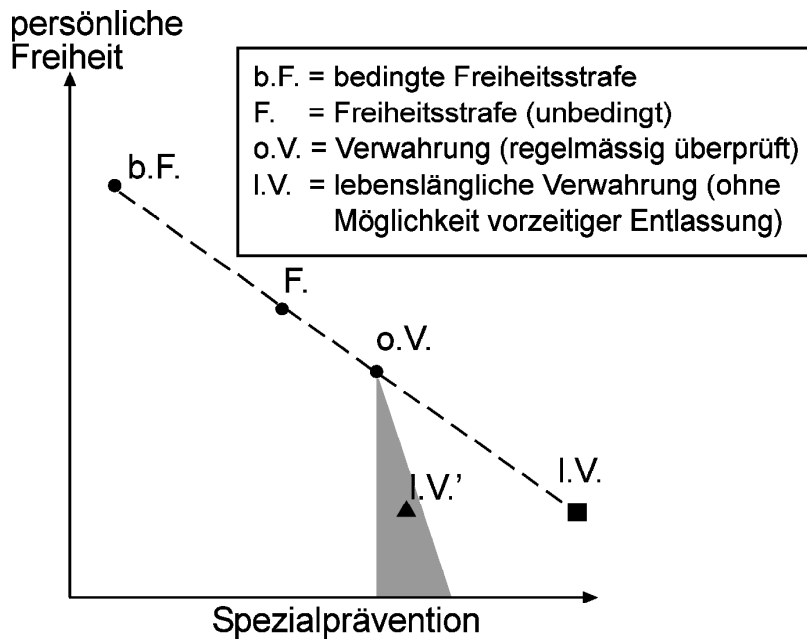


Abbildung 21: Trade-off zwischen persönlicher Freiheit und Spezialprävention

- 770 Die bedingte Freiheitsstrafe bewirkt die geringste Einschränkung der persönlichen Freiheit, hat aber auch die geringste spezialpräventive Wirkung. Damit will ich nicht behaupten, dass die Rückfallgefahr nach einer vollzogenen Freiheitsstrafe geringer ist als bei einer bedingten Strafe. Aber während des Vollzugs der Freiheitsstrafe kann der Täter nicht weiter delinquieren, so dass die Spezialprävention zumindest während dieser Zeit sichergestellt ist. Die unbedingte Freiheitsstrafe stellt daher die zweite Stufe dar; d.h. grössere Einschränkung der persönlichen Freiheit, aber auch grössere spezialpräventive Wirkung. Die ordentliche Verwahrung mit regelmässiger Überprüfung nach Art. 64 Abs. 1 Vorentwurf stellt die dritte Stufe dar; sie ist auf jeden Fall eine stärkere Einschränkung der persönlichen Freiheit als die unbedingte Freiheitsstrafe, weil der Vollzug der Freiheitsstrafe der Verwahrung voraus geht (der Täter somit länger eingesperrt bleibt), aber sie hat auch eine grössere spezialpräventive Wirkung.
- 771 Bis zur (vorgeschlagenen) Einführung der lebenslänglichen Verwahrung war die ordentliche Verwahrung gleichzeitig *ultima ratio*, die extremste Wahl. Mit der lebenslänglichen Verwahrung steht den Gerichten nun eine zusätzliche, extremere Alternative zur Auswahl. Ob man die lebenslängliche Verwahrung eher beim Viereck oder beim Dreieck in der Abbildung 21 ansiedeln will, hängt davon ab, welche Vorteile man sich von ihr verspricht. Die Befürworter der lebenslänglichen Verwahrung werden sie eher beim Viereck, d.h. in der linearen Fortsetzung der Trendlinie, ansiedeln: auch sie werden kaum bestreiten, dass die lebenslängliche Verwahrung eine stärkere Beschränkung der persönlichen Freiheit bedeutet als die ordentliche Verwahrung. Sie machen aber geltend, dass dieser Nachteil durch die grössere spezialpräventive Wirkung aufgewogen wird. Gegner der lebenslänglichen Verwahrung sind der Auffassung, dass der Verlust an persönlicher Freiheit (und Rechtsstaatlichkeit) durch den geringen zusätzlichen Gewinn an öffentlicher Sicherheit nicht kompensiert wird. Sie werden die lebenslängliche Verwahrung eher dort ansiedeln, wo das Dreieck in Abbildung 21 ist, also im grauen Bereich der klar minderwertigen Alternativen.

- 772 Für die folgende Analyse ist es irrelevant, ob die lebenslängliche Verwahrung im Bereich der Trendlinie oder bei den klar minderwertigen Alternativen angesiedelt wird: Kontrast- und Kompromisseffekt legen auf jeden Fall nahe, dass *mit der Einführung der lebenslänglichen Verwahrung der Anteil der ordentlichen Verwahrungen zu Lasten der „gewöhnlichen“ Freiheitsstrafen (ohne Verwahrung) steigen wird*. Ich habe diese Voraussage empirisch getestet.

V. Eigene Studie zum Kompromisseffekt

A. Frage zum Kompromisseffekt

- 773 Den Richterinnen und Richter der Zivil- und Strafgerichte der Kantone beider Basel, Bern und Graubünden wurde folgender Sachverhalt zur Beurteilung vorgelegt (2004):

X ist seit 1993 drei Mal unter anderem wegen strafbarer Handlungen gegen die sexuelle Integrität zu Gefängnisstrafen verurteilt worden (1993 zu 6 Monaten, 1997 zu 8 Monaten, 2000 zu 10 Monaten). Meist handelte es sich um Exhibitionismus, wobei er einmal auch ein kleines Mädchen ausgegriffen hatte.

Nach seiner Entlassung aus dem Strafvollzug Anfang November 2002 zog X bei einer ihm bekannten allein erziehenden Mutter zweier Knaben (Jahrgang 1993 und 1996) als Untermieter ein und übernahm während der Abwesenheit der Mutter Betreuungsfunktionen. Während der Abwesenheit der Mutter zeigte X dem jüngeren Knaben mehrfach seinen erigierten Penis. Teilweise befriedigte sich X vor den Augen des jüngeren Knaben selber; einmal beobachtete der ältere Knabe dies. X versuchte, den jüngeren Knaben zum Oralsex zu zwingen, sah aber davon ab, als dieser sich wehrte.

X schrieb 2003 in einem Primarschulhaus an fünf Orten seine Telefonnummer auf, mit einem pornographischen Text und der Aufforderung, "ruft mich an, für Mädchen von 7 - 19 Jahren".

- 774 Es handelt sich bei diesem Sachverhalt um den aufgrund der Ergebnisse des Pretests angepassten Sachverhalt, der dem Bundesgerichtsurteil 6S.782/2000 vom 20. Dezember 2000 zugrunde lag (nicht in der amtlichen Sammlung publiziert). Gemäss dem Sachverhalt, der dem Bundesgerichtsurteil zugrunde lag, hatte der Täter versucht, mit dem jüngeren Knaben Analsex zu vollziehen, den er abbrach, als der Knabe sich über Schmerzen beklagte. Auch waren seine Vorstrafen weit zahlreicher als im obigen Sachverhalt. Beide kantonalen Vorinstanzen hatten den Verurteilten verwahrt. Das Bundesgericht hob den Entscheid der Vorinstanz auf und wies den Fall zur erneuten Beurteilung an das Kantonsgericht zurück. Es kritisierte, das vorinstanzliche Urteil beantworte die Frage nicht, „ob sich die Straftaten des Beschwerdeführers gegenüber dem jüngeren Knaben als einmalig oder als Höhepunkt einer stetigen Verschlimmerung des sexuellen Krankheitsbildes darstellen“ (Urteil 6S.782/2000, E. 1d). Daher könne ohne Ergänzung des Sachverhalts nicht entschieden werden, ob eine schwerwiegende Gefährdung der öffentlichen Sicherheit im Sinne von Art. 43 Ziff. 1 Abs. 2 StGB vorliege. Da sich diese Frage kaum objektiv beantworten lässt, schien das Urteil gut geeignet als Stimulus-Material, da vernünftige Menschen geteilter Meinung sein können, ob eine Verwahrung angezeigt ist.
- 775 Der Pretest ergab jedoch, dass ausnahmslos alle Versuchspersonen den Täter verwahrten. Damit war die Fragestellung nicht geeignet, nachzuweisen, dass die Einführung der zusätzlichen, extremeren, Alternative „lebenslängliche Verwahrung“ den Anteil der Verwahrung gegenüber der Freiheitsstrafe steigern kann. Wenn der „Marktanteil“ der Verwahrung

bereits 100 % ist, kann er sich nicht mehr steigern. Für die Umfrage im Herbst 2004 wurde der Sachverhalt daher „entschärft“, d.h. der versuchte Analverkehr und einige weitere geringere Delikte wurden weggelassen und die Anzahl der Vorstrafen reduziert. Damit sollte eine gleichmässige Verteilung zwischen Freiheitsstrafe und Verwahrung erreicht werden. Aus dem gleichen Grund ist die Freiheitsstrafe relativ hoch angesetzt; so erscheint sie eher als valable Alternative zur Verwahrung. Möglicherweise sind die Anpassungen etwas zu weit gegangen; einige Richter bemerkten jedenfalls – nicht ganz zu Unrecht – dass eine Verwahrung bei diesem Sachverhalt wohl nicht ernsthaft in Frage komme. Die Manipulation hat aber, wie die Resultate zeigen, funktioniert.

- 776 Den Richtern in der Gruppe „ohne lebenslängliche Verwahrung“ wurde folgende Frage gestellt:

Welche der zur Auswahl stehenden Strafen/Massnahmen betrachten Sie in diesem Fall als angemessen?
Die lebenslängliche Verwahrung ohne Möglichkeit der vorzeitigen Entlassung steht nicht zur Verfügung.

- ☐ Drei Jahre Gefängnis (unbedingt) mit begleitender ambulanter Psychotherapie.
- ☐ Drei Jahre Gefängnis, wobei der Vollzug der Strafe zugunsten einer Verwahrung auf unbestimmte Zeit aufgeschoben wird. Der Angeklagte wird entlassen, wenn feststeht, dass er keine Gefahr für die Gesellschaft mehr darstellt.

- 777 Der Hinweis darauf, dass die lebenslängliche Verwahrung nicht zur Verfügung steht, wurde eingefügt, um sicherzustellen, dass die unterschiedlichen Anteile der verschiedenen Optionen nicht auf einen unterschiedlichen Informationsstand zurückzuführen sein können. Wenn die Option lebenslängliche Verwahrung zur Verfügung steht, kann dies dazu führen, dass man sich Gedanken macht, welche Aspekte des Sachverhalts *für* die lebenslängliche Verwahrung sprechen. Damit wird der Blick auf die massnahmeschärfenden Merkmale des Sachverhalts gelenkt, was eine Verschiebung von der Freiheitsstrafe zur ordentlichen Verwahrung erklären kann.¹¹⁹⁵ Da die Richterinnen und Richter in der Gruppe ohne lebenslängliche Verwahrung ausdrücklich auf die lebenslängliche Verwahrung aufmerksam gemacht wurden, kann diese Erklärung ausgeschlossen werden.

- 778 Den Richterinnen und Richter in der Gruppe „mit lebenslänglicher Verwahrung“ wurde der gleiche Sachverhalt und die gleiche Frage vorgelegt; mit dem Unterschied, dass die lebenslängliche Verwahrung als weitere Option bei den möglichen Strafen und Massnahmen aufgeführt war (und natürlich der Hinweis, dass die lebenslängliche Verwahrung nicht zur Verfügung steht, fehlte). Die Gruppe „lebenslängliche Verwahrung“ hatte also neben den oben erwähnten Möglichkeiten der Verurteilung die folgende zusätzliche Option:

- ☐ Drei Jahre Gefängnis, wobei der Vollzug der Strafe zugunsten einer lebenslänglichen Verwahrung ohne Möglichkeit der vorzeitigen Entlassung aufgeschoben wird.

B. Resultate

- 779 Von den 234 Richterinnen und Richter, die den Fragebogen zumindest teilweise ausgefüllt zurückgeschickt haben, haben 230 die Frage nach der angemessenen Strafe oder Mass-

¹¹⁹⁵ Zu diesen Bedenken generell KELMAN/ROTTENSTREICH/TVERSKY, FN 1185, 71 f.

nahme beantwortet. Die Resultate sind in der Tabelle 25 dargestellt. In der Gruppe ohne Möglichkeit der lebenslänglichen Verwahrung wird die Freiheitsstrafe klar präferiert. Rund zwei Drittel der Richterinnen und Richter wählen die Freiheitsstrafe. In der Gruppe mit der Möglichkeit der lebenslänglichen Verwahrung sinkt der Anteil der Freiheitsstrafen auf 45 %, während die ordentlichen Verwahrungen auf 53 % steigen. Da der „Marktanteil“ der ordentlichen Verwahrung bei der Hinzufügung einer weiteren Alternative steigt, wird die Regularität verletzt. Die lebenslängliche Verwahrung – die aufgrund des Sachverhalts nicht angezeigt ist – wird nur zwei Mal gewählt. Der Unterschied zwischen den beiden Gruppen ist statistisch signifikant.¹¹⁹⁶

	Gruppe ohne	Gruppe mit	Total
Gefängnis	78 (65 %)	50 (45 %)	128
Ordentl. Verwahrung	41 (34 %)	59 (53 %)	100
Lebenslängl. Verwahrung	–	2 (2 %)	2
Total	119	111	230

Tabelle 25: Resultate der Frage zum Kompromisseffekt

C. Diskussion

- 780 Die Resultate bestätigen die Ergebnisse von KELMAN et al. Durch das Hinzufügen einer extremen Option zum *choice set* kann der Anteil der mittleren Option gesteigert werden. Die Verschiebung um rund 20 Prozentpunkte ist vergleichbar mit dem Ausmass des von KELMAN et al. beobachteten Kompromisseffekts bei juristischen Entscheidungen; sie berichten von Verschiebungen (je nach Sachverhalt) von 15-20 Prozentpunkten.¹¹⁹⁷ Inwiefern sich die Resultate in die Wirklichkeit übertragen lassen, ist eine offene Frage. Aufgrund der zahlreichen Studien zum Kompromiss- und Kontrasteffekt kann man zwar davon ausgehen, dass der Effekt robust ist. Andererseits wurde das *choice set* hier – wie in den meisten anderen Studien – künstlich eingeschränkt. In Wirklichkeit hat der Richter oft mehr als zwei oder drei Entscheidungsmöglichkeiten, die zudem nicht klar abgegrenzt sind. Ob und wie Kontrast- und Kompromisseffekte dann wirken, ist eine offene Frage, obwohl die Resultate von KIVETZ et al. darauf hindeuten, dass der Kompromisseffekt auch in komplexen Wahlsituationen seine Wirkung zeigt.¹¹⁹⁸ Dass die Effekte vollständig verschwinden, ist nicht anzunehmen und wird durch die Studien, die unter realistischeren Bedingungen durchgeführt wurden, nicht nahe gelegt.¹¹⁹⁹
- 781 Die Resultate, kombiniert mit denen anderer Studien, machen zumindest glaubhaft, dass Kompromiss- und Kontrasteffekte juristische Entscheidungen zu beeinflussen vermögen. Anekdotische Evidenz stützt diese Annahme. Jeremy Stephenson, Präsident des Strafgerichts Basel-Stadt, äusserte sich gegenüber der SonntagsZeitung dahingehend, dass er kaum je eine lebenslängliche Verwahrung aussprechen werde und sich stattdessen für die

¹¹⁹⁶ $\chi^2 = 11,100$; $p = 0,004$.

¹¹⁹⁷ KELMAN/ROTTENSTREICH/TVERSKY, FN 1185, 64, 67.

¹¹⁹⁸ KIVETZ/NETZER/SRINIVASAN, FN 1177, 252.

¹¹⁹⁹ KELMAN/ROTTENSTREICH/TVERSKY, FN 1185, 73.

ordentliche Verwahrung entscheiden würde.¹²⁰⁰ Damit hat er natürlich nicht gesagt, dass er sich inskünftig *häufiger* für die ordentliche Verwahrung entscheiden wird; das soll ihm auch nicht unterstellt werden. Aber die ordentliche Verwahrung lässt sich leichter rechtfertigen, wenn daneben noch die lebenslängliche Verwahrung zur Verfügung steht – man hat den Täter dann ja „nur“ ordentlich verwahrt, nicht lebenslänglich. Dieser Effekt führt zu einer Verschiebung der Urteile zu Gunsten der ordentlichen Verwahrung.

- 782 Ob man den Einfluss von Kontrast- und Kompromisseffekt durch explizite Warnungen verringern kann, wurde meines Wissens bisher nicht empirisch untersucht. Da der Grossteil der Studien zum Kontrast- und Kompromisseffekt aus Untersuchungen zum Konsumentenverhalten stammen, ist dies nicht erstaunlich. Marketing-Leute sind in erster Linie interessiert zu erfahren, unter welchen Umständen die Effekte am stärksten wirken, und weniger daran, wie man sie bekämpfen kann.

1. Erkenntnisse für den Gesetzgeber

- 783 Aus normativer Sicht ist die Kontextabhängigkeit juristischer Urteile bedenklicher als die Kontextabhängigkeit der Entscheidungen von Konsumenten. Die Entscheidung eines Konsumenten für oder gegen ein Produkt kann kaum je als falsch oder richtig bezeichnet werden, weil es keinen normativen Standard für die richtige Wahl gibt. Die Wahl hängt eben von den Präferenzen des Konsumenten ab, und diese ändern sich offenbar je nachdem, wie viele Produkte zur Auswahl stehen. Von einem „Fehler“ des Konsumenten kann man allenfalls dann sprechen, wenn er seinen Entscheid nach der Aufklärung über den Kontrast- und Kompromisseffekt bereut und rückgängig machen möchte.¹²⁰¹
- 784 Die Kontextabhängigkeit juristischer Urteile ist problematisch, weil juristische Entscheide erstens nicht den Entscheidenden, sondern Dritte treffen, und zweitens durch (mehr oder weniger) explizite Normen bestimmt sein sollten. Der Entscheid, jemanden zu verwahren, sollte davon abhängen, ob von ihm eine schwerwiegende Gefährdung der öffentlichen Sicherheit ausgeht, die nicht durch andere Massnahmen verhindert werden kann (Art. 59, 64 Abs. 1 Vorentwurf). Der Entscheid sollte nicht davon beeinflusst werden, ob zudem die Möglichkeit der lebenslänglichen Verwahrung besteht, wenn der Täter eine besonders schwere Anlasstat begangen hat und „im Vergleich zu anderen Tätern solcher Delikte eine sehr hohe Wahrscheinlichkeit dafür besteht, dass er ein weiteres Verbrechen dieser Art begeht, und der Täter aufgrund besonderer Persönlichkeitsmerkmale als dauerhaft nicht therapierbar eingestuft wird“ (Art. 64 Abs. 1^{ter} lit. a und b Vorentwurf). Der Gesetzgeber muss sich bewusst sein, dass er durch die Einführung einer weiteren, scheinbar klar abgegrenzten, Kategorie nicht nur die Fälle regelt, die in diese Kategorie fallen, sondern sich auch das Verhältnis der bisherigen Kategorien verschiebt – ohne dass die formellen Abgrenzungskriterien verändert würden. Dies gilt auch, wenn die neue Kategorie kaum je angewandt wird. Dies mag beim neuen Institut der lebenslänglichen Verwahrung der Fall

¹²⁰⁰ SonntagsZeitung vom 11. Juli 2004, 5.

¹²⁰¹ KELMAN/ROTTENSTREICH/TVERSKY, FN 1185, 74.

sein,¹²⁰² aber ändert nichts daran, dass sich der Anteil der ordentlichen Verwahrungen erhöhen wird.

2. Erkenntnisse für den Praktiker

- 785 Was Immobilienagenten und Gebrauchtwagenhändlern recht ist, kann dem Anwalt billig sein. Wer die Wirkung von Kontrast- und Kompromisseffekt kennt, kann sie leicht zu Gunsten seines Klienten einsetzen. Ein Beispiel von KELMAN et al. zur Illustration: Herr W. ist Nachbar eines Nachtclubs und fühlt sich durch den Lärm gestört. Eine Klage gegen den Nachtclub hätte mit grosser Wahrscheinlichkeit Erfolg und würde dazu führen, dass der Club seinen Betrieb erheblich einschränken müsste. Der Betreiber des Clubs macht Herrn W. daher folgendes Angebot: er würde entweder schalldämmende Massnahmen ergreifen oder Herrn W. jedes Wochenende die Unterkunft in einem guten Hotel sowie \$ 120 bar für seine Umstände bezahlen. Einer zweiten Gruppe macht der Betreiber das gleiche Angebot, aber zusätzlich offeriert er auch noch, die Übernachtung in einem guten Hotel zu bezahlen und Herrn W. zusätzlich jede Woche \$ 40 bar und \$ 85 in nicht übertragbaren Gutscheinen, die in verschiedenen Nachtclubs eingelöst werden können, zu schenken. Die Versuchspersonen wurden darauf aufmerksam gemacht, dass Herr W. im Schnitt nur drei Mal pro Jahr in einen Nachtclub geht und dort in der Regel weniger als \$ 90 ausgibt.
- 786 Die Option „Übernachtung und Gutscheine“ ist daher klar minderwertig verglichen mit der Option „Übernachtung und Bargeld“. Die Versuchspersonen wurden gefragt, welches der Vergleichsangebote sie Herrn W. zur Annahme empfehlen würden. Die Gruppe, denen die klar minderwertige Option nicht zur Verfügung stand, empfahl zu 47 % die Übernachtung mit Bargeld und zu 53 % die schalldämmenden Massnahmen. Die Gruppe, die unter den drei Alternativen wählte, empfahl zu 74 % die Übernachtung mit Bargeld und zu 26 % die schalldämmenden Massnahmen (niemand empfahl die klar minderwertige Alternative).¹²⁰³ Ein Anwalt, der die Gegenseite auf die Option „Übernachtung mit Bargeld“ hinsteuern will, tut daher gut daran, zusätzlich eine ähnliche, aber in einer Dimension minderwertige Option anzubieten (und zwar von Anfang an; die nachträgliche Ergänzung des *choice set* ist, wie die Resultate von HUBER et al. beim *within-subject* Design zeigen, weniger effektiv).¹²⁰⁴
- 787 Bei der Formulierung von Eventualbegehren in Zivilprozessen muss man sich bewusst sein, dass die mittlere Option attraktiver erscheint, als sie im direkten Vergleich zum ersten oder dritten Begehren erscheinen würde. Die entsprechenden taktischen Überlegungen liegen auf der Hand.

¹²⁰² LAURENT MOREILLON in der SonntagsZeitung vom 19. September 2004, 5.

¹²⁰³ KELMAN/ROTTENSTREICH/TVERSKY, FN 1185, 71.

¹²⁰⁴ HUBER/PAYNE/PUTO, FN 1146, 94.

§ 17 Selbstüberschätzung und *overconfidence*

I. Selbstüberschätzung

- 788 Menschen neigen dazu, sich und ihre Fähigkeiten egozentrisch und eigennützig zu beurteilen.¹²⁰⁵ Eigene Erfolge werden Können zugeschrieben, Misserfolge äusseren Umständen. Sie schätzen sich selber regelmässig als überdurchschnittlich ein in Bezug auf begehrtenwerte Eigenschaften wie Gesundheit,¹²⁰⁶ berufliche Fähigkeiten,¹²⁰⁷ Autofahren¹²⁰⁸ und der Wahrscheinlichkeit, eine glückliche Ehe zu führen.¹²⁰⁹ Leute überschätzen ihren Beitrag zu gemeinsamen Aktivitäten. Beispielsweise nehmen beide Parteien nach einem Gespräch an, mehr als die Hälfte der Zeit gesprochen zu haben.¹²¹⁰ Ebenso addieren sich die Anteile, die jeder Partner nach eigener Einschätzung zur Haushaltsarbeit beigetragen hat, in der Regel auf mehr als 100 %.¹²¹¹ Und natürlich unterliegen nur die anderen kognitiven Täuschungen.¹²¹²
- 789 Selbstüberschätzung hat verschiedene Ursachen. Erstens sehen sich die meisten Menschen selber als kompetent und sachkundig. Eine Überschätzung der eigenen Fähigkeiten stimmt mit diesem Selbstbild überein.¹²¹³ Zweitens wollen sich die meisten Menschen in einem guten Licht darstellen.¹²¹⁴ Sie glauben vielleicht nicht tatsächlich, dass sie besser als der Durchschnitt sind, aber geben dies gegenüber Dritten an.¹²¹⁵ Drittens suchen sie in erster Linie nach Informationen, die ihre Annahme – beispielsweise, dass sie sich nicht scheiden lassen werden – unterstützen (siehe vorne, S. 181 ff.). Da sie keine vergleichbaren Informationen über die Ehen Dritter haben, schliessen sie beispielsweise aus den vorhandenen Daten, dass ihre Ehe überdurchschnittlich ist.¹²¹⁶ Viertens ist Erinnerung eigennützig in dem Sinn, dass eigene Handlungen besser erinnert werden als fremde Taten.¹²¹⁷ Wenn wir über unseren Anteil an der gemeinsamen Hausarbeit gefragt werden, erinnern wir uns

¹²⁰⁵ MICHAEL ROSS/FIORE SICOLY, Egocentric Biases in Availability and Attribution, in: KAHNEMAN/SLOVIC/TVERSKY (Hrsg.), 179-189.

¹²⁰⁶ NEIL D. WEINSTEIN, Unrealistic Optimism About Future Life Events, Journal of Personality and Social Psychology 1980, 806-820, 809 ff.

¹²⁰⁷ K. PATRICIA CROSS, Not Can, But Will College Teaching Be Improved, New Directions for Higher Education 1997, 1-15, 9 f.

¹²⁰⁸ OLA SVENSON, Are We All Less Risky and More Skillful Than Our Fellow Drivers?, Acta Psychologica 1980, 143-148, 145 f.

¹²⁰⁹ BAKER/EMERY, FN 191, 441 ff.: trotz Kenntnis der Scheidungsrate von 50 % (in den USA, in der Schweiz beträgt sie ca. 43 %) erwarten praktische alle frisch Verheirateten, dass ihre Ehe nicht geschieden wird.

¹²¹⁰ ROSS/SICOLAY, FN 1205, 182.

¹²¹¹ ROSS/SICOLAY, FN 1205, 183 f.

¹²¹² EMILY PRONIN/DANIEL Y. LIN/LEE ROSS, The Bias Blind Spot: Perceptions of Bias in Self Versus Others, Personality and Social Psychology Bulletin 2002, 369-381.

¹²¹³ HART BLANTON/BRETT W. PELHAM/TRACY DEHART/MAURICIO CARVALLO, Overconfidence as Dissonance Reduction, Journal of Experimental Social Psychology 2001, 373-385.

¹²¹⁴ SUSAN T. FISKE/SHELLEY E. TAYLOR, Social Cognition, New York 1991, 78 ff.

¹²¹⁵ FISKE/TAYLOR, FN 1214, 78 ff.

¹²¹⁶ BAKER/EMERY, FN 1209, 446 f.

¹²¹⁷ FISKE/TAYLOR, FN 1214, 78 ff.

besser an die Arbeiten, die wir erledigt haben, und überschätzen daher deren Anteil an der Gesamtarbeit.¹²¹⁸ Schliesslich sind zahlreiche positive Eigenschaften unklar definiert, und jeder der Befragten erfüllt daher möglicherweise seine persönliche Definition der Eigenschaft.¹²¹⁹ Beispielsweise verstehen Personen unter einem guten Autofahrer ganz Verschiedenes, und daher kann sich jeder einbilden, ein überdurchschnittlicher Autofahrer zu sein – immer gemäss seinen persönlichen Kriterien.¹²²⁰

II. Overconfidence

- 790 In der Forschung zum Urteilen unter Unsicherheit ist der Begriff *overconfidence* ein Terminus technicus, der besagt, dass Leute zu grosses Vertrauen in ihr Urteil haben.¹²²¹ In typischen Studien zum *overconfidence* Phänomen werden Leute nach dem Wahrheitsgehalt einer Aussage gefragt („Winterthur hat mehr Einwohner als St. Gallen: Stimmt/Stimmt nicht“) und gleichzeitig gebeten, anzugeben, wie sicher sie sich sind, dass ihre Antwort richtig ist, ausgedrückt auf einer Skala von 50 % („könnte auch eine Münze werfen,“) bis 100 % („todsicher“).¹²²² Die so ausgedrückte subjektive Sicherheit wird als „Konfidenz“ bezeichnet.¹²²³ Die Resultate werden typischerweise in einer Kalibrierungskurve gemäss Abbildung 22 dargestellt; d.h. man misst die relative Häufigkeit richtiger Antworten („Richtigkeit“, *accuracy*) auf alle Fragen, für die die Versuchspersonen eine bestimmte Konfidenz angegeben haben, und trägt diesen Wert auf die Konfidenz ab. Die Diagonale stellt die Norm für die richtige Kalibrierung dar – einer Konfidenz von 60 % entspricht eine relative Häufigkeit von 60 % richtigen Antworten, einer Konfidenz von 80 % eine relative Häufigkeit von 80 % richtigen Antworten u. s. w. Die Kalibrierungskurve liegt aber typischerweise unter dieser Diagonalen; d.h. die Verlässlichkeit des eigenen Wissens wird überschätzt (dies gilt für die aggregierten Daten; die Kalibrierungskurve lässt nicht den Schluss zu, dass jeder einzelne Versuchsteilnehmer schlecht kalibriert ist). Je schwieriger die Frage – d.h. je seltener die richtige Antwort – desto grösser die *overconfidence* („unmöglich“ ist eine Frage, die im Schnitt nicht häufiger als 50 % richtig beantwortet wird). Bei sehr einfachen Fragen, die von den meisten Versuchspersonen richtig beantwortet werden, tritt hingegen das umgekehrte Phänomen der *underconfidence* auf, d.h. die Leute sind sich ihres Wissens zu wenig sicher. Dies wird als *hard-easy effect* bezeichnet.¹²²⁴ Die Kalibrierung von Experten ist derjenigen von Laien typischerweise nur überle-

¹²¹⁸ ROSS/SICOLY, FN 1205, 324.

¹²¹⁹ LOEWENSTEIN/ISSACHAROFF/CAMERER/BABCOCK, FN 910, 141.

¹²²⁰ SVENSON, FN 1208, 145.

¹²²¹ ULRICH HOFFRAGE, Die Illusion der Sicherheit bei Entscheidungen unter Unsicherheit, in: HELL/FIEDLER/GIGERENZER (Hrsg.), FN 11, 73-97, 73.

¹²²² Z.B. SARAH LICHTENSTEIN/BARUCH FISCHHOFF, Do Those Who Know More Also Know More about How Much They Know?, *Organizational Behavior and Human Performance* 1977, 159-183. Winterthur hat übrigens mehr Einwohner als St. Gallen.

¹²²³ HOFFRAGE, FN 1221, 74.

¹²²⁴ SARAH LICHTENSTEIN/BARUCH FISCHHOFF/LAWRENCE D. PHILLIPS, Calibration of Probabilities: The State of the Art to 1980, in: KAHNEMAN/SLOVIC/TVERSKY (Hrsg.), 306-334, 316.

gen, wenn der Experte eine schnelle, unmissverständliche Rückmeldung erhält, ob seine Voraussage richtig gewesen ist – wie dies beispielsweise bei Meteorologen der Fall ist.¹²²⁵

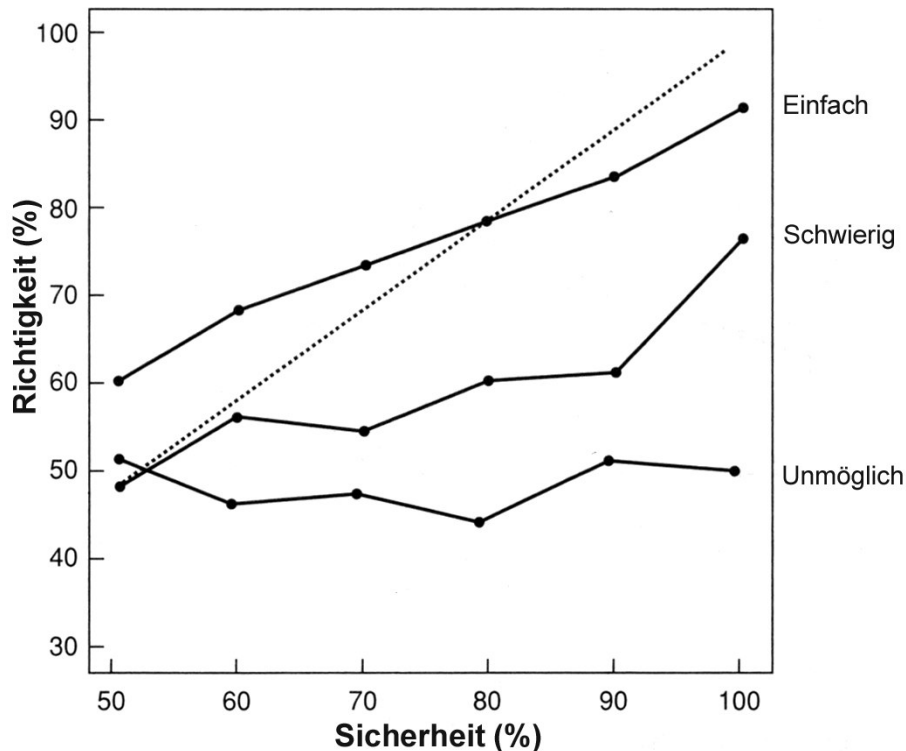


Abbildung 22: Typische Kalibrierungskurven (aus GRIFFIN/TVERSKY, 1992/2002)

- 791 Je sicherer sich die Versuchspersonen sind, desto grösser ist die *overconfidence*.¹²²⁶ Gerade wenn sich die Probanden zu 100 % sicher sind, dass ihre Antwort stimmt, ist die *overconfidence* (gemessen als Konfidenz minus relative Häufigkeit richtiger Antworten) besonders gross.¹²²⁷ In den „todsicheren“ Fällen sind in der Regel nur 85-90 % der Antworten richtig. Dies gilt insbesondere auch für Augenzeugen – der Zeuge, der sich zu 100 % sicher ist, dass er den Angeklagten am Tatort gesehen hat, irrt sich also in einem von zehn Fällen.¹²²⁸
- 792 Die Forschung zum *overconfidence* Phänomen, das während zwanzig Jahren als robust gelten hatte, ist Anfang der 90-er Jahre von zwei Seiten angegriffen worden. Die einen Kritiker führen an, dass der beobachtete Effekt ein statistisches Artefakt ist, die anderen, dass er verschwindet, wenn die Fragen zufällig ausgewählt werden.

¹²²⁵ LICHTENSTEIN/FISCHHOFF/PHILLIPS, FN 1224, 306.

¹²²⁶ JOSHUA KLAYMAN/JACK S. BELL/CLAUDIA GONZALEZ-VALLEJO/SEMA BARLAS, Overconfidence: It Depends on How, What, and Whom You Ask, *Organization Behavior and Human Decision Processes* 1999, 216-247, 237; ELIZABETH F. LOFTUS/WILLEM A. WAGENAAR, Lawyers' Predictions of Success, *Jurimetrics* 1987, 437-453, 444.

¹²²⁷ BARUCH FISCHHOFF/PAUL SLOVIC/SARAH LICHTENSTEIN, Knowing with Certainty: The Appropriateness of Extreme Confidence, *Journal of Experimental Psychology: Human Perception and Performance* 1977, 552-564.

¹²²⁸ KENNETH A. DEFFENBACHER, Eyewitness Accuracy and Confidence: Can we Infer Anything about their Relationship?, *Law and Human Behavior* 1980, 243-260.

- 793 Das statistische Argument ist, dass sich die richtige Antwort auf die gestellte Frage aus zwei Komponenten zusammensetzt: der Leistung der Antwortenden und einem zufälligen Element. Betrachtet man die Richtigkeit als Funktion der Konfidenz, d.h. sagt man die Anzahl richtiger Ergebnisse aufgrund des Konfidenz-Levels voraus, so liegen die aufgrund der Konfidenz vorausgesagten Werte wegen des zufälligen Elements und der daraus folgenden nicht perfekten Regression näher beim Mittelwert – m. a. W. beobachtet man eine Regression zum Mittelwert.¹²²⁹ Die beobachteten Abweichungen sind nach dieser Ansicht nicht auf einen systematischen Fehler zurückzuführen, sondern sind bei jedem zufälligen Fehler zu erwarten und hängen vom Ausmass des Fehlers resp. der Korrelation zwischen Konfidenz und richtigen Antworten ab. Wenn das Ausmass des zufälligen Fehlers bekannt ist – weil die Versuchspersonen die gleiche Frage mehrmals beantworten und die Unterschiede gemessen werden – kann man die aufgrund des zufälligen Fehlers erwartete *overconfidence* berechnen und mit der beobachteten vergleichen. So lässt sich feststellen, ob die beobachtete *overconfidence* nur auf den unsystematischen Fehler zurückzuführen ist. Zwei neue Studien, die diese Methode verwenden, kommen unabhängig voneinander zum Schluss, dass systematische *overconfidence* existiert und nicht ausschliesslich durch zufällige Fehler erklärt werden kann; allerdings fällt sie geringer aus, als bisher angenommen.¹²³⁰
- 794 Der zweite Einwand gegen die traditionelle Forschung zum *overconfidence* Phänomen ist, dass die Versuchsleiter scheinbar einfache Fragen verwenden, die aber schwierig sind.¹²³¹ Man könnte dies den „Fangfragen-Vorwurf“ nennen. Nach GIGERENZER et al. konstruiert die Versuchsperson ein Probabilistisches Mentales Modell, wenn sie mit einer Frage wie „Welche Stadt hat mehr Einwohner, Münster oder Bremen?“ konfrontiert ist. Wenn die Versuchspersonen kein sicheres Wissen über die Einwohnerzahl hat, versucht sie, aus ihr bekannten Hinweisen (*cues*) auf die Einwohnerzahl zu schliessen. Die Hinweise „hat Flughafen“ oder „hat Bundesliga-Verein“ deuten darauf hin, dass die eine Stadt eine gewisse Grösse hat. Wenn sie nur bei einer der beiden Städte vorliegen, hat diese Stadt wahrscheinlich mehr Einwohner. Wenn der Versuchsleiter nun absichtlich ein Städtepaar für seine Frage auswählt, bei der nur die kleinere Stadt einen Bundesliga-Verein hat, werden die Versuchspersonen in die Irre geführt, weil ein in der ökologischen Realität valider Hinweis plötzlich wertlos ist. Nach dem Modell von GIGERENZER et al. müsste daher die *overconfidence* weitgehend verschwinden, wenn den Versuchspersonen repräsentative, zufällig ausgewählte Fragen gestellt werden; in der eigenen Untersuchung von GIGE-

¹²²⁹ ROBYN M. DAWES/MATT MULFORD, The False Consensus Effect and Overconfidence: Flaws in Judgment or Flaws in How We Study Them?, *Organizational Behavior and Human Decision Making Processes* 1996, 201-211; IDO EREV/THOMAS S. WALLSTEN/DAVID V. BUDESCU, Simultaneous Over- and Underconfidence: The Role of Error in Judgment Processes, *Psychological Review* 1994, 519-527.

¹²³⁰ DAVID V. BUDESCU/THOMAS S. WALLSTEN/WING TUNG WU, On the Importance of Random Error in the Study of Probability Judgment. Part II: Applying the Stochastic Judgment Model to Detect Systematic Trends, *Journal of Behavioral Decision Making* 1997, 173-188; KLAYMAN/BELL/GONZALEZ-VALLEJO/BARLAS, FN 1227, 217 ff.

¹²³¹ GERD GIGERENZER/ULRICH HOFFRAGE/HANS KLEINBÖLTING, Probabilistic Mental Models: A Brunswikian Theory of Confidence, *Psychological Review* 1991, 506-528; PETER JUSLIN, The Overconfidence Phenomenon as a Consequence of Informal Experimenter Guided Selection of Almanac Items, *Organizational Behavior and Human Decision Processes* 1994, 226-246.

RENZER et al. ist dies auch der Fall.¹²³² Verschiedene unabhängige Studien konnten diese Ergebnisse aber nicht replizieren.¹²³³ GRIFFIN und KAHNEMAN weisen zudem darauf hin, dass die von GIGERENZER verwendeten Städtevergleichs-Fragen einfacher waren als die Allgemeinwissensfragen der Kontrollgruppe (72 % richtige gegenüber 53 % richtigen Antworten) und daher auch der bekannte *hard-easy effect* das Verschwinden der *overconfidence* in den Studien von GIGERENZER et al. erklären kann.¹²³⁴

- 795 Wenn der Urteilende in Bezug auf Aufgaben, die er aus seiner Erfahrungswelt kennt, gut kalibriert ist, sollten Experten wie Psychiater,¹²³⁵ Ärzte,¹²³⁶ Investment-Banker,¹²³⁷ Unternehmer,¹²³⁸ Manager,¹²³⁹ Anwälte,¹²⁴⁰ Busfahrer und Fussballer¹²⁴¹ ihre Fähigkeit, auf dem Gebiet ihrer jeweiligen Expertise richtig zu urteilen oder zu handeln (beispielsweise den Kursverlauf von Aktien vorherzusagen¹²⁴², Hirnschädigungen¹²⁴³ oder Krebs¹²⁴⁴ korrekt zu diagnostizieren oder mit einem Bus zwischen zwei Pfosten durchfahren, ohne diese zu berühren)¹²⁴⁵ nicht systematisch überschätzen. Die zitierten Studien zeigen, dass sie dies aber regelmässig tun.¹²⁴⁶
- 796 Abbildung 23 zeigt die Kalibrierung von amerikanischen und holländischen Anwälten in tatsächlichen Gerichtsfällen.¹²⁴⁷ Sowohl amerikanische wie holländische Anwälte zeigen

¹²³² GIGERENZER/HOFFRAGE/KLEINBÖLTING, FN 1231, 525.

¹²³³ LYLE A. BRENNER/DEREK J. KOEHLER/VARDA LIBERMAN/AMOS TVERSKY, Overconfidence in Probability and Frequency Judgments: A Critical Examination, *Organizational Behavior and Human Decision Processes* 1996, 212-219; zahlreiche weitere Nachweise bei KLAYMAN/BELL/GONZALEZ-VALLEJO/BARLAS, FN 1227, 220.

¹²³⁴ GRIFFIN/TVERSKY, FN 906, 244.

¹²³⁵ STUART OSKAMP, Overconfidence in Case-Study Judgments, *Journal of Consulting Psychology* 1965, 261-265; Nachdruck in: KAHNEMAN/SLOVIC/ TVERSKY (Hrsg.), 287-293.

¹²³⁶ JAY J. CHRISTENSEN-SZALANSKI/JAMES B. BUSHYHEAD, Physicians' Use of Probabilistic Information in Real Clinical Setting, *Journal of Experimental Psychology: Human Perception and Performance* 1981, 928-935.

¹²³⁷ CARL-AXEL S. STAËL VON HOLSTEIN, Probabilistic Forecasting: An Experiment Related to Stock Market, *Organizational Behavior and Human Performance* 1972, 139-158.

¹²³⁸ ARNOLD C. COOPER/CAROLYN Y. WOO/WILLIAM C. DUNKELBERG, Entrepreneurs' Perceived Chances for Success, *Journal of Business Venturing* 1988, 97-108.

¹²³⁹ J. EDWARD RUSSO/PAUL J.H. SHOEMAKER, Managing Overconfidence, *Sloan Management Review* 1992, 7-17.

¹²⁴⁰ LOFTUS/WAGENAAR, FN 1226.

¹²⁴¹ Nachweise bei NIGEL HARVEY, Relations Between Confidence and Skilled Performance, in: WRIGHT/AYTON (Hrsg.), FN 619, 321-352, 331 ff.

¹²⁴² GUSTAF TORNGREN/HENRY MONTGOMERY, Worse Than Chance? Performance and Confidence Among Professionals and Laypeople in the Stock Market, *Journal of Behavioral Finance* 2004, 148-153.

¹²⁴³ LEWIS R. GOLDBERG, The Effectiveness of Clinicians' Judgments: The Diagnosis of Organic Brain Damage from the Bender-Gestalt Test, *Journal of Consulting Psychology* 1959, 25-33.

¹²⁴⁴ CHRISTENSEN-SZALANSKI/BUSHYHEAD, FN 1236, 930.

¹²⁴⁵ HARVEY, FN 1241, 331.

¹²⁴⁶ Siehe auch KOEHLER/BRENNER/GRIFFIN, FN 392.

¹²⁴⁷ Die Daten stammen von MARIJKE MALSCH, Lawyers' Predictions of Judicial Decisions, Dissertation Leiden 1989; und JANE GOODMAN-DELAHUNTY/PAR-ANDERS GRANHAG/ELIZABETH F. LOFTUS, How Well Can Lawyers Predicts Their Chances of Success?, U. of Washington 1998 (nicht publiziert); die Grafik aus KOEHLER/BRENNER/GRIFFIN, FN 392, 706.

overconfidence, wenn sie sich relativ sicher sind, den Fall zu gewinnen, und *underconfidence*, wenn sie mit grosser Wahrscheinlichkeit mit einer Niederlage rechnen. Die holländischen Anwälte sind besser als die amerikanischen Anwälte in der Lage, zwischen erfolgreichen und nicht erfolgreichen Fällen zu unterscheiden. Je näher der Verhandlungstermin rückt – und je mehr sie sich folglich mit dem Fall beschäftigt haben – desto besser wird die Kalibrierung der amerikanischen Anwälte. Wenn der Erfolg persönlich relevant ist, weil die Anwälte auf der Basis eines Erfolgshonorars arbeiten, ist die *overconfidence* ausgeprägter: Anwälte, die auf Erfolgsbasis arbeiteten, erwarten im Schnitt wie die Anwälte, die nach Stunden abrechnen, in 65 % der Fälle zu gewinnen, gewinnen aber sehr viel weniger häufig (42 % der Fälle verglichen mit 56 %).¹²⁴⁸

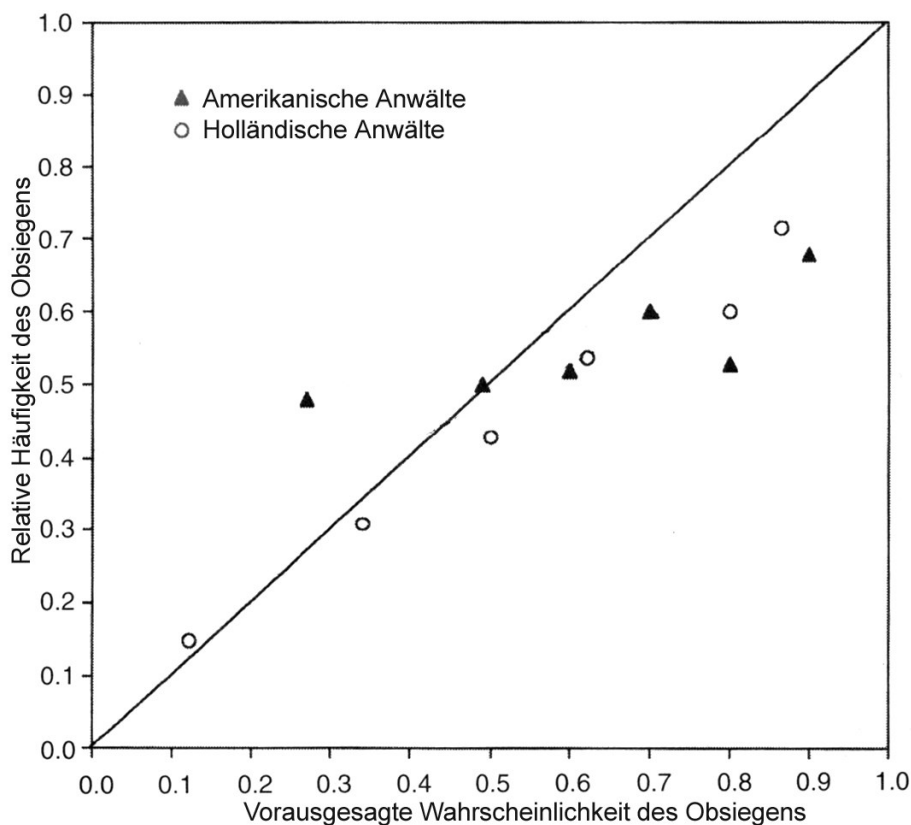


Abbildung 23: Kalibrierung amerikanischer und holländischer Anwälte (aus KOEHLER et al., 2002)

- 797 Dass Geschworene sich ihrer Entscheidungen manchmal allzu sicher sind, steht fest. HUGO BEDAU und MICHAEL RADELET berichten von 350 amerikanischen Urteilen, in denen jemand erwiesenermassen zu Unrecht wegen eines Kapitalverbrechens verurteilt wurde – obwohl die Geschworenen der Ansicht waren, die Schuld stehe „*beyond reasonable doubt*“ fest.¹²⁴⁹ Allerdings verlangt das Recht auch bei Straftaten, die mit der Todesstrafe bedroht sind, nicht, dass die Schuld mit Sicherheit feststeht. Bei den 350 Fällen handelt es sich

¹²⁴⁸ KOEHLER/BRENNER/GRIFFIN, FN 392, 706.

¹²⁴⁹ HUGO BEDAU/MICHAEL RADELET, Miscarriages of Justice in Potentially Capital Cases, Stanford Law Review 1987, 21-90. Fünf Fälle wurden durch Berufungsgerichte korrigiert, 67 Personen wurden zu Gefängnis bis zu 25 Jahren, 139 zu lebenslänglichem Gefängnis verurteilt, 139 zum Tod. 23 waren bereits hingerichtet worden, als der Artikel erschien.

daher möglicherweise nur um diejenigen Fehltritte, die man bewusst in Kauf nimmt, wenn man verlangt, dass die Schuld bloss „mit an Sicherheit grenzender Wahrscheinlichkeit“ feststehen muss.

III. Kulturelle und geschlechtliche Unterschiede bei Selbstüberschätzung und *overconfidence*

- 798 Bei Selbstüberschätzung und *overconfidence* existieren kulturelle Unterschiede, die sich allerdings nicht immer in der Richtung bewegen, die man gemeinhin erwarten würde. So erwarten die meisten Menschen, dass Asiaten weniger zu allzu grossem Vertrauen in das eigene Urteil neigen wie grossmäulige Amerikaner. Tatsächlich ist es aber so, dass chinesische, malaysische und indonesische Studenten mehr *overconfident* sind als amerikanische.¹²⁵⁰ Bei der Einschätzung bezüglich karriererelevanter Eigenschaften – wie z.B. der Fähigkeit, sich schriftlich ausdrücken zu können – stufen sich taiwanesischen, indischen und amerikanischen Studenten als überdurchschnittlich, japanische und singapurische Studenten hingegen bescheiden als leicht unterdurchschnittlich ein.¹²⁵¹ Es besteht kein zwingender Zusammenhang zwischen Selbsteinschätzung und *overconfidence*; so sind die bescheidenen Singapurer sich in ihrem Urteil allzu sicher, während die ebenfalls bescheidenen Japaner gut kalibriert sind; die Inder sind bescheidener als die Amerikaner, sich ihres Urteils aber (fälschlicherweise) viel sicherer als diese.¹²⁵² LEE et al. schreiben die *overconfidence* teilweise dem Schulsystem zu; ihrer Auffassung nach führen Schulsysteme, die grossen Wert auf Auswendiglernen und Frontalunterricht legen, zu mehr *overconfidence* als Schulsysteme, die Wert auf Verstehen und Diskussion legen, weil sich die Absolventen der dogmatischen Systeme nicht gewohnt sind, Argumente pro und kontra vorzustellen.¹²⁵³
- 799 Bei den Geschlechtern hingegen ist der Unterschied so, wie ihn wohl die meisten erwarten: Männer neigen mehr zu Selbstüberschätzung und *overconfidence* als Frauen – und bezahlen dafür auch, in dem sie beispielsweise ihr Vermögen erwiesenermassen suboptimal verwalten.¹²⁵⁴ Männer neigen vor allem bezüglich als „maskulin“ wahrgenommener Fähigkeiten zur Selbstüberschätzung (wozu Spekulieren am Aktienmarkt und Autofahren gehören).¹²⁵⁵

¹²⁵⁰ J. FRANK YATES/JU-WHEI LEE/WINSTON R. SIECK/INCHEOL CHOI/PAUL C. PRICE, Probability Judgment across Cultures, in: GILOVICH/GRIFFIN/KAHNEMAN (Hrsg.), 271-291.

¹²⁵¹ JU-WHEI LEE/J. FRANK YATES/HIROMI SHINOTSUKA/RAMADHAR SINGH/MARY LOU UY ONGLATCO/NAI-SHENG YEN/MEENAKSHI GUPTA/DEEPTI BHATNAGAR, Cross-National Differences in Overconfidence, Asian Journal of Psychology 1995, 63-69, 66.

¹²⁵² LEE/YATES/SHINOTSUKA/SINGH/UY ONGLATCO/YEN/GUPTA/BHATNAGAR, FN 1251, 66.

¹²⁵³ LEE/YATES/SHINOTSUKA/SINGH/UY ONGLATCO/YEN/GUPTA/BHATNAGAR, FN 1251, 67.

¹²⁵⁴ BRAD M. BARBER/TERRANCE ODEAN, Boys Will Be Boys: Gender, Overconfidence, And Common Stock Investment, The Quarterly Journal of Economics 2001, 261-292.

¹²⁵⁵ BARBER/ODEAN, FN 1254, 264

IV. Overconfidence und Konfliktlösung

- 800 Es leuchtet unmittelbar ein, dass *overconfidence* die effiziente Lösung von Konflikten verhindern kann.¹²⁵⁶ Wenn beide Parteien überzeugt sind, zu gewinnen, werden sie sich nicht auf einen Vergleich einigen können. LOEWENSTEIN und Kollegen baten Jura-Studierende, die Erfolgsaussichten einer Klage zu schätzen. Ein Teil der Testpersonen spielte die Rolle der Kläger, ein anderer Teil diejenige der Beklagten. Beide Teilgruppen erhielten die gleichen Informationen über den Fall.¹²⁵⁷ Als sie gebeten wurden, vorauszusagen, wie hoch der durch das Gericht zugesprochene Schadenersatz ausfallen würde, sagten die Studierenden, die den Fall aus Sicht des Klägers zu beurteilen hatten, \$ 14'527 mehr voraus als die Studierenden, die den Fall aus beklagterischer Sicht beurteilten. Die Kläger beurteilten ein Vergleichsangebot als fair, das \$ 17'709 über demjenigen lag, das die Beklagten als fair empfanden.¹²⁵⁸ Die eigennützige Wahrnehmung der Fakten führt dazu, dass ein Vergleich nicht zustande kommt, weil beide Parteien überzeugt sind, vor Gericht besser abzuschneiden. Da die verzerrte Wahrnehmung bei selbstrelevanten Informationen stärker ist, kann es eine wichtige Aufgabe des Anwalts sein, die Perspektive des neutralen Dritten einzunehmen, um seinen Klienten vor allzu überzogenen Erwartungen abzubringen.

V. Eigene Studie zur Selbsteinschätzung

A. Frage zur Selbsteinschätzung

- 801 Die Richterinnen und Richter der Kantone Aargau, St. Gallen und Zürich wurden gebeten, sich selber mit ihren Kollegen und Kolleginnen zu vergleichen:

Die Qualität eines Richters oder einer Richterin misst sich unter anderem am Anteil seiner/ihrer Entscheide, gegen die erfolgreich Berufung oder Rekurs erhoben wird (ob dieses Kriterium gerechtfertigt ist, bleibe dahingestellt).

Angenommen, es gäbe eine Rangliste aller schweizerischen Richter und Richterinnen. Zuoberst auf der Liste stünden die RichterInnen, deren Entscheide am seltensten durch eine obere gerichtliche Instanz korrigiert werden.

Wo stünden Sie selber auf dieser Liste?

- 802 Zur Auswahl standen vier Möglichkeiten, nämlich „ich würde zu den besten 25% gehören“, „ich würde zu den besseren 50% gehören“, „ich würde zu den schlechteren 50% gehören“ und „ich würde zu den schlechtesten 25% gehören“.

B. Resultate

- 803 156 der 181 Richterinnen und Richter, die den Fragebogen der ersten Umfrage zurückgeschickt hatten, beantworteten die Frage nach der Selbsteinschätzung. Damit zeigte diese

¹²⁵⁶ LINDA BABCOCK/GEORGE LOEWENSTEIN, Explaining Bargaining Impasse: The Role of Self-Serving Biases, *Journal of Economic Perspectives* 1997, 109-126.

¹²⁵⁷ LOEWENSTEIN/ISSACHAROFF/CAMERER/BABCOCK, FN 910, 145.

¹²⁵⁸ LOEWENSTEIN/ISSACHAROFF/CAMERER/BABCOCK, FN 910, 150.

Frage die höchste Ausfallrate der Umfrage 2003. Viele Richter empfanden es offenbar als unangenehm, sich selber beurteilen zu müssen. Einige bemerkten, sie könnten die Frage nicht beantworten, da sie immer im Gremium entscheiden würden.

- 804 Von den 156 Richterinnen und Richtern, die die Frage beantworteten, stuften sich 20 % (31) im besten Quartil ein, 72 % (112) im zweitbesten Quartil, 7 % (11) im zweitschlechtesten Quartil und 1,3 % (2) im schlechtesten Quartil. 92 % der Richter stuften sich selber also als überdurchschnittlich ein – was offensichtlich nicht zutreffen kann.

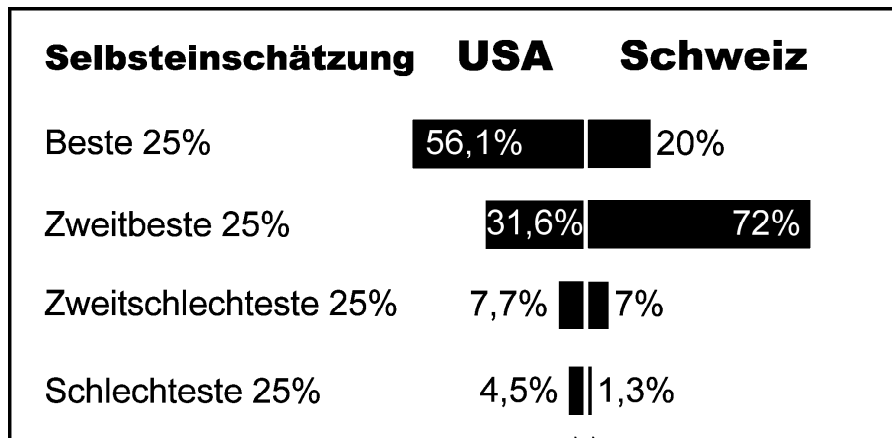


Abbildung 24: Selbstüberschätzung: Vergleich Schweiz-USA

- 805 Interessant ist der Vergleich mit den amerikanischen Richtern der Studie von GUTHRIE/RACHLINSKI/WISTRICH.¹²⁵⁹ Beinahe gleich viele amerikanische wie schweizerische Richter stufen sich als überdurchschnittlich ein (87,7 % resp. 92 %). Die schweizerischen Richter stufen sich selber jedoch seltener bei den Besten ein. Sie unterschätzen die eigene Stellung diesbezüglich sogar, denn nur 20 % stufen sich bei den besten 25 % ein.

C. Diskussion

- 806 In der Schweiz gibt es fast ausschliesslich überdurchschnittliche Richter, wenn man der Selbsteinschätzung Glauben schenkt. Ob die Richter selber überzeugt sind von ihren Antworten, oder diese nur gaben, um gegenüber Dritten gut dazustehen, lässt sich aufgrund der Resultate nicht beurteilen. Da die Richter wussten, dass die Umfrage anonym war, bestand für sie wenig Anlass, ihre wirkliche Selbsteinschätzung zu verbergen.
- 807 Resultate ähnlicher Studien zeigen, dass Menschen tatsächlich meist glauben, überdurchschnittlich zu sein. Der Wunsch, sich selber in einem besseren Licht darzustellen, vermag nur einen Teil der Resultate zu erklären.¹²⁶⁰ Es ist daher wahrscheinlich, dass die Richter (und hier sind die Richterinnen für einmal nicht mitgemeint) ihre eigenen Fähigkeiten tatsächlich überschätzen.

¹²⁵⁹ GUTHRIE/RACHLINSKI/WISTRICH, FN 20, 814.

¹²⁶⁰ BABCOCK/LOEWENSTEIN, FN 1256, 110 f.

- 808 Bezüglich des Ausmasses der Selbstüberschätzung scheint es einen wichtigen kulturellen Unterschied zu geben. Die Schweizer haben offenbar Hemmungen, sich selber als zu den Besten gehörend zu bezeichnen – selbst in einer anonymen Befragung. Man ist lieber „im oberen Mittelfeld“ – geradezu unterdurchschnittlich will man natürlich auch wieder nicht sein.
- 809 Selbstüberschätzung kann einen Richter dazu verleiten, dort nicht zu zweifeln, wo er zweifeln sollte. Der Richter beurteilt möglicherweise die Folge eines Verhaltens als vorhersehbar oder vermeidbar, die es objektiv nicht war; er lehnt unter Umständen ein Ausstandsbegehren ab, weil er nicht an seiner Fähigkeit zweifelt, den Fall objektiv und unparteiisch zu beurteilen. Generell gesagt macht es Selbstüberschätzung schwieriger für Richter und Richterinnen, einzusehen und zuzugeben, dass sie Fehler machen.
- 810 Andererseits ist ein leicht übertriebener Glaube an die eigenen Fähigkeiten wichtig für die psychische Gesundheit.¹²⁶¹ Zumindest in westlichen Kulturen scheint richtige Selbsteinschätzung mit leichten Depressionen einherzugehen.¹²⁶² Selbstüberschätzung führt auch zu Selbstsicherheit, und die meisten Leute ziehen einen selbstsicheren, entscheidungsfreudigen Richter einem ängstlichen, unsicheren vor. Unter dem Strich profitiert das Rechtssystem daher vielleicht von der Tendenz der Richter, sich selber zu überschätzen.

¹²⁶¹ SHELLEY E. TAYLOR/JONATHAN D. BROWN, Illusion and Well-Being: A Social Psychological Perspective on Mental Health, *Psychological Bulletin* 1988, 193-210.

¹²⁶² SHELLEY E. TAYLOR, *Positive Illusions: Creative Self-Deception and the Healthy Mind*, New York 1989, 214.

§ 18 Zusammenfassung/Thesen

1. Die amerikanischen Rechtsrealisten bezweifelten, dass Gerichtsurteile ausschliesslich durch Rechtsregeln bestimmt werden. Sie suchten daher nach anderen Gründen, juristische Urteile zu erklären und auch zu rechtfertigen, und wandten sich dazu insbesondere den Sozialwissenschaften und der Psychologie zu. Ihr Erbe ist das instrumentelle, d.h. auf Folgeerwägungen abstellende, Rechtsverständnis der amerikanischen Juristen.
2. Die Ökonomische Analyse des Rechts ist die Anwendung der Ökonomik, d.h. der wirtschaftswissenschaftliche Methode, auf rechtliche Fragestellungen. Sie versucht einerseits das Verhalten von Rechtsunterworfenen, Rechtsstab und Gesetzgeber durch ökonomische Gesetze zu erklären (positive ÖAR) und misst andererseits Rechtsnormen daran, ob sie das wohlfahrtsökonomische Effizienzziel verwirklichen (normative ÖAR). Kritisiert wird sowohl der normative Status der Effizienz als Rechtsziel als auch das Menschenbild des rationalen Nutzenmaximierers (*homo oeconomicus*), das der Ökonomik zu Grunde liegt.
3. Seit den frühen siebziger Jahren des letzten Jahrhunderts wurden in der psychologischen Forschung zahlreiche Verletzungen der Axiome der Erwartungsnutzentheorie (*expected utility theory*) festgestellt. Diese systematischen und vorhersehbaren Verletzungen („Fehler“) führen dazu, dass die Voraussagen von ökonomischen Modellen, die auf den Axiomen der Erwartungsnutzentheorie beruhen, nachweislich falsch sind.
4. In der Wirtschaftswissenschaft wird seit den frühen achtziger Jahren unter dem Stichwort „*behavioral economics*“ versucht, diese empirischen Erkenntnisse zu verwenden. In der Rechtswissenschaft haben amerikanische Juristen in den neunziger Jahren begonnen, die ÖAR auf Basis der psychologischen Forschung zu kritisieren und die Erkenntnisse der Psychologie für die Analyse juristischer Fragestellungen zu nutzen. Die Strömung ist unter der Bezeichnung „*behavioral law and economics*“ bekannt. Nach der hier vertretenen Auffassung wird sie dadurch unnötigerweise auf die Kritik an der traditionellen Ökonomischen Analyse des Rechts reduziert. Besser betrachtet man den Versuch, psychologische Erkenntnisse auf breiter Front im Recht anzuwenden, als eine neue Form der Rechtspsychologie, die sich nicht auf einzelne Akteure des Rechtssystems (wie Geschworene, Augenzeugen, Kriminelle) beschränkt, sondern alle Rechtsunterworfenen, den Rechtsstab und den Gesetzgeber in ihr Blickfeld einbezieht. So verstanden ist die Strömung eine neue Art des Rechtsrealismus auf empirischer Grundlage.
5. Im zweiten Teil der Arbeit wird der Einfluss einzelner so genannter „kognitiver Täuschungen“ auf Entscheidungen von Richtern und Parteien in Zivil- und Strafverfahren untersucht. Es wurden zwei briefliche Umfragen durchgeführt, an denen im Herbst 2003 und 2004 insgesamt 415 Richterinnen und Richter der Kantone Aargau, beider Basel, Bern, Graubünden, St. Gallen und Zürich teilnahmen.
6. Als *Ankereffekt* (*anchoring*) wird die Assimilation eines numerischen Urteils an einen vorgegebenen Vergleichsstandard bezeichnet. Der Vergleichsstandard wirkt

wie ein „Anker“ für die endgültige Schätzung und zieht diese in seine Richtung. Der Einfluss des Ankereffekts ist besonders stark, wenn wenige Informationen über den richtigen Zielwert vorhanden sind. Regeln, die den Ermessensspielraum des Gerichts einschränken, sind geeignet, dem Einfluss des Ankereffekts entgegenzuwirken. In Verhandlungssituationen führt der Ankereffekt dazu, dass diejenige Partei, die das erste Angebot macht und daher den ersten Anker setzen kann, im Vorteil ist.

7. Als *Darstellungseffekt (framing)* wird die Tatsache bezeichnet, dass Menschen sich in der gleichen Entscheidungssituation unterschiedlich verhalten, je nachdem, ob die Optionen als Gewinne oder Verluste dargestellt werden. Personen, die eine Wahl als Alternative zwischen Verlusten wahrnehmen, verhalten sich risikogeneigt. Das kann im Gegensatz zu traditionellen ökonomischen Modellen erklären, warum Beklagte rund 70-80 % aller Prozesse verlieren: weil sie sich risikogeneigt verhalten, lehnen sie selbst für sie günstige Vergleichsvorschläge ab. Beklagte sollten daher in ihrem eigenen Interesse vermehrt ermutigt werden, einen Vergleich abzuschliessen. Allerdings verhalten sich auch Richter unbewusst risikogeneigter, wenn sie einen Vergleichsvorschlag aus der Sicht der Beklagten beurteilen; d.h. sie empfehlen häufiger, den Vergleich abzulehnen.
8. Als *Unterlassungseffekt (omission bias)* wird die Tatsache bezeichnet, dass Menschen die negativen Folgen einer Unterlassung in der Regel als weniger moralisch vorwerfbar als die gleichen Folgen einer Handlung empfinden. Der Unterlassungseffekt ist möglicherweise darauf zurückzuführen, dass Unterlassung oft (aber eben nicht immer) fahrlässig sind, während Handlungen absichtlich begangen werden. Der Unterlassungseffekt kann die logisch nur schwer zu rechtfertigende Unterscheidung zwischen aktiver und passiver Sterbehilfe erklären.
9. Intuitives statistisches Denken führt oft zu Fehlern. In Gerichtsverfahren sollte daher vermehrt das *Bayes-Theorem*, das die kohärente Integration neuer Indizien mit vorbestehenden Überzeugungen erlaubt, zumindest als Urteilsheuristik verwendet werden. D.h. es ist nicht gefordert, dass Urteile mit mathematischer Scheingenauigkeit „errechnet“ werden, sondern dass die Urteilsbegründung den Gesetzen der Wahrscheinlichkeitstheorie zumindest verbal Rechnung trägt. Dadurch lassen sich zahlreiche Denkfehler bei der Beweiswürdigung vermeiden.
10. Bei der Beweiswürdigung von DNA-Gutachten wird oft übersehen, dass die Beweiskraft eines DNA-Gutachtens entscheidend davon abhängt, wie oft das Labor zu einem *falsch positiven* Resultat gelangt. Als falsch positiv wird ein „Treffer“ bezeichnet, der tatsächlich keiner ist; er kann beispielsweise auf eine Kontamination oder Verwechslung der Proben zurückzuführen sein. Empirische Daten aus den USA legen nahe, dass die Möglichkeit falsch positiver DNA-Gutachten nicht rein theoretisch ist. In der Schweiz werden bisher keine Tests durchgeführt, die es erlauben würden, die falsch positiv Rate akkreditierter forensischer DNA-Labors abzuschätzen.
11. Unter *Bestätigungstendenz (confirmation bias)* wird das Phänomen verstanden, dass Informationen, die eine von der Versuchsperson zumindest in Betracht gezogene Hypothese bestätigen, von ihr eher gesucht, wahrgenommen, stärker gewichtet oder

besser in Erinnerung behalten werden als Informationen, die gegen die Hypothese sprechen. Die Bestätigungstendenz ist besonders stark, wenn man motiviert ist, eine Hypothese zu verteidigen. Juristen sollten sich daher nicht frühzeitig auf eine Hypothese festlegen, schon gar nicht öffentlich. Anwälte müssen sich bewusst sein, dass ihre Sicht des Falles oft durch ihre einseitige Perspektive als Interessenvertreter verzerrt ist.

12. Als *Regression zum Mittelwert* wird das statistische Gesetz bezeichnet, dass ein Extremwert, der auch durch zufällige Ursachen zu Stande gekommen ist, bei einer erneuten Messung zum Mittelwert hin tendiert. Da juristische Interventionen wie beispielsweise Strafen und Massnahmen meist in dem Moment erfolgen, in dem ein Extremwert (der Delinquenz) erreicht wird, wird der Einfluss der Intervention auf das weitere Verhalten des Betroffenen überschätzt. Die beobachtete Besserung ist zumindest vermutungsweise oft auf die Regression zum Mittelwert und nicht auf die Wirkung der Strafe oder Massnahme zurückzuführen. Das fehlende Verständnis der Regression zum Mittelwert kann erklären, warum viele Richter trotz gegenteiliger kriminologischer Erkenntnisse von der positiven Wirkung von harten Strafen überzeugt sind.
13. Als *Rückschaufehler (hindsight bias)* wird die Tendenz bezeichnet, die Voraussehbarkeit eines Ereignisses im nachhinein zu überschätzen („das musste ja schief gehen“). Der Rückschaufehler ist höchstwahrscheinlich eine Nebenfolge adaptiven Lernens. Gerichte müssen jedoch bei Haftpflichtprozessen und bei Strafverfahren wegen Fahrlässigkeitsdelikten eine *ex ante* Perspektive einnehmen und beurteilen, ob die Folgen einer Handlung oder Unterlassung voraussehbar waren. Der Rückschaufehler legt es nahe, dass sie dabei die Voraussehbarkeit oft überschätzen, da im Zeitpunkt des Urteils die Folgen natürlich bekannt sind.
14. Unter dem *Hofeffekt (halo effect)* wird die Tendenz verstanden, faktisch unabhängige oder nur mässig korrelierende Eigenschaften von Personen oder Sachen fälschlicherweise als zusammenhängend wahrzunehmen. Als „Schönheits-Hofeffekt“ wird die Tatsache bezeichnet, dass attraktive Menschen häufig als erfolgreicher, intelligenter und besser eingeschätzt werden als ihre weniger attraktiven Mitmenschen. Ob die Attraktivität des Angeklagten einen Einfluss auf die Strafzumessung hat, wurde in zahlreichen Experimenten und in mindestens zwei Feldstudien untersucht, deren Ergebnisse teilweise widersprüchlich sind. Die vorliegende Arbeit konnte keinen Effekt der Attraktivität der Angeklagten auf die Strafzumessung nachweisen.
15. Als *Kontrasteffekt* wird die Tatsache bezeichnet, dass eine Alternative in einer Wahlsituation an Attraktivität gewinnt, wenn daneben eine asymmetrisch dominierte Alternative zur Auswahl steht. Als *Kompromisseffekt (auch extremeness aversion)* bezeichnet man das Phänomen, dass die extremen Alternativen in einem *Choice Set* zu Gunsten der mittleren („Kompromiss“) Optionen vernachlässigt werden. Durch die Änderung der zur Auswahl stehenden Alternativen können sich daher die Präferenzen für die Alternativen verändern. Präferenzen sind mit anderen Worten *kontextabhängig*. Dies kann beispielsweise Folgen haben, wenn ein neuer Straftatbestand oder eine neue Rechtsfolge eingeführt wird. Am Beispiel der lebenslänglichen Verwahrung ohne Möglichkeit der vorzeitigen Entlassung, die in der Schweiz aufgrund

Besonderer Teil

einer Volksinitiative eingeführt werden soll, wird gezeigt, dass sich dadurch die Präferenz für die „normale“ Verwahrung, die von der extremen Option zur Kompromissoption wird, verändert.

16. Als *Overconfidence* wird die Überschätzung der eigenen Fähigkeiten oder des eigenen Wissens bezeichnet. Nachweislich werden von Parteien und Anwälten die Aussichten der eigenen Sache vor Gericht systematisch überschätzt. Overconfidence kann dazu führen, dass Vergleichsverhandlungen rationaler Parteien scheitern.

Anhang A: Herleitung des Bayes-Theorems aus den Axiomen der Wahrscheinlichkeitstheorie

811 In den folgenden Grafiken bezeichnet S den Ereignisraum, die Menge aller möglichen Ereignisse. Ein Ereignis ist eine Teilmenge von S .

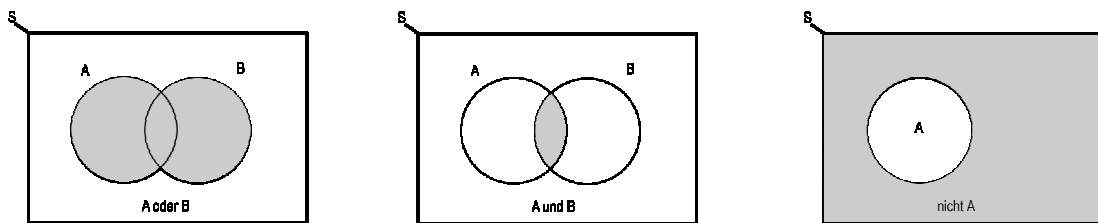
812 Zwei Ereignisse A und B können in folgenden Kombinationen auftreten

- $A \cup B$ (A oder B), d.h. es treten entweder A oder B oder beide zusammen auf (*Gesamtmenge* der beiden Ereignisse);
- $A \cap B$ (A und B); d.h. es tritt sowohl A als auch B auf (*Schnittmenge* der beiden Ereignisse). Wenn A und B sich gegenseitig ausschliessen, ist die Schnittmenge leer.

813 Weiter bedeutet

- $\neg A$ *Ausschluss* des Ereignisses A und bezeichnet alle Ereignisse, die nicht A sind.

814 Grafisch lassen sich $A \cup B$, $A \cap B$ und $\neg A$ wie folgt darstellen:¹²⁶³



815 Die klassischen Axiome der Wahrscheinlichkeitstheorie sind

- a) Die Wahrscheinlichkeit eines Ereignisses hat einen Wert zwischen 0 und 1, wobei 0 bedeutet, dass das Ereignis nie eintritt und 1, dass es sicher eintritt.
- b) Die Randwahrscheinlichkeit aller Ereignisse im Ereignisraum S ist 1.
- c) Wenn zwei Ereignisse A und B sich gegenseitig ausschliessen gilt $A \cup B = P(A) + P(B)$ und $P(A \cap B) = 0$. Dies gilt für alle n sich gegenseitig ausschliessender Ereignisse. Beispiel: Wenn ich mit einer Wahrscheinlichkeit von 0,5 zur Nachmittagsvorlesung gehe und mit einer Wahrscheinlichkeit von 0,3 ins Seebad, dann beträgt die Wahrscheinlichkeit, dass ich entweder zur Nachmittagsvorlesung oder ins Seebad gehe, 0,8. Die Wahrscheinlichkeit, dass ich sowohl ins Seebad als auch in die Nachmittagsvorlesung gehe, ist 0.

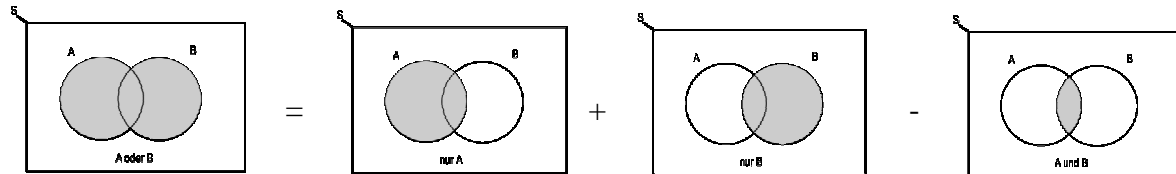
¹²⁶³ Die Darstellung und die Herleitung folgt CHAMPOD/TARONI, FN 710, 214 ff. Für den formalen Beweis siehe JEFFREY, FN 613, 16.

Besonderer Teil

- 816 Wenn sich die beiden Ereignisse A und B nicht gegenseitig ausschliessen, berechnet sich die Wahrscheinlichkeit, dass entweder A oder B oder beide zusammen auftreten, nach der folgenden Regel:

$$A \cup B = P(A) + P(B) - A \cap B$$

- 817 Grafisch dargestellt:



- 818 Durch die Addition von A und B wird der Bereich, wo sich A und B überschneiden, zwei Mal gezählt. Er muss daher wieder einmal subtrahiert werden.
- 819 Mit ganz ähnlichen Überlegungen lässt sich herleiten, wie die bedingte Wahrscheinlichkeit $P(B | A)$, d.h. die Wahrscheinlichkeit, dass B auftritt, wenn wir wissen, dass A vorliegt, aus den Wahrscheinlichkeiten $P(A \cap B)$ und $P(A)$ berechnet werden kann:
- 820 Wenn A und B sich gegenseitig ausschliessen, liegt die Antwort auf der Hand: B wird nicht eintreten. Wenn A und B sich nicht gegenseitig ausschliessen, ist es möglich, dass B eintritt. Da wir wissen, dass A eingetreten ist, reduziert sich der Ereignisraum S auf die Teilmenge A. Was uns interessiert sind die Ereignisse B, die innerhalb der Teilmenge A liegen, oder mit anderen Worten die Schnittmenge $A \cap B$. Die bedingte Wahrscheinlichkeit $P(B | A)$ ergibt sich aus der Division von $P(A \cap B)$ durch $P(A)$, d.h. aus dem Verhältnis der Fläche „A und B“ zur Fläche „nur A“ gemäss den obigen Grafiken. Es gilt also

$$P(B | A) = \frac{P(A \cap B)}{P(A)} \quad \text{oder} \quad P(A \cap B) = P(B | A) \cdot P(A)$$

- 821 Das Bayes-Theorem ergibt sich aus der Überlegung, dass $P(A \cap B) = P(B \cap A)$, oder
- $$P(B | A) \cdot P(A) = P(A | B) \cdot P(B)$$

- 822 Nach $P(B | A)$ gelöst

$$P(B | A) = \frac{P(A \cap B) \cdot P(B)}{P(A)}$$

- 823 $P(A)$ lässt sich zerlegen in

$$P(A) = \underbrace{P(B \cap A)}_{P(A | B) \cdot P(B)} + \underbrace{P(\neg B \cap A)}_{P(A | \neg B) \cdot P(\neg B)} = P(A | B) \cdot P(B) + P(A | \neg B) \cdot P(\neg B)$$

824 Daher lautet das Bayes-Theorem in seiner extensiven Form, für zwei sich gegenseitig ausschliessende Ereignisse A und B

$$P(B | A) = \frac{P(B) \cdot P(A | B)}{P(B) \cdot P(A | B) + P(\neg B) \cdot P(A | \neg B)}$$

Anhang B: Fragebogen 2003

825 *Jede der Fragen begann auf einer neuen Seite. Die Varianten sind mit [VERSION 1/2/3] gekennzeichnet.*

826 Titelblatt: Fragebogen für Richterinnen und Richter

Dieser Fragebogen soll helfen, Erkenntnisse über einige Aspekte richterlicher Entscheidungsfindung zu gewinnen. Die Resultate der Umfrage werden statistisch ausgewertet und in einem wissenschaftlichen Aufsatz publiziert. Die Umfrage erfolgt absolut anonym. Die Beantwortung der Fragen sollte nicht mehr als zehn Minuten in Anspruch nehmen. Es ist sehr wichtig, dass Sie diesen Fragebogen selbständig, ohne die Fragen vor der Beantwortung mit Kollegen und Kolleginnen zu besprechen, beantworten. Bitte senden Sie den Fragebogen bis Ende dieser Woche mit dem beigelegten Rückantwortkuvert zurück. Für Ihre Mitarbeit danken wir Ihnen ganz herzlich.

827 Frage 1: Höhe der Genugtuung

Das Auto der Klägerin wurde von einem Lastwagen erfasst. Der Lastwagen hatte der Klägerin den Rechtsvortritt verweigert. Der Lastwagenfahrer war im Zeitpunkt des Unfalles betrunken. Seine Blutalkoholkonzentration betrug 1,6 Promille. Die Klägerin, eine 20-jährige Frau, fuhr korrekt und konnte den Unfall nicht vermeiden.

Die Klägerin, die vor dem Unfall in guter körperlicher Verfassung war, ist durch die unmittelbaren Folgen des Unfalles querschnittgelähmt. Eine Heilung ist ausgeschlossen. Nach rund dreimonatigem Aufenthalt im Spital konnte sie nach Hause entlassen werden. Sie wird ihren erlernten Beruf als Detailhandelsverkäuferin nicht mehr ausüben können und wird auch auf ihr Hobby, das Reiten, verzichten müssen.

Die Klägerin hat sich mit der Haftpflichtversicherung des Halters des Lastwagens aussergerichtlich über die Höhe des Schadenersatzes geeinigt. Die beklagte Versicherung anerkennt auch den Anspruch der Klägerin auf Genugtuung im Grundsatz. Streitig ist in diesem Zeitpunkt einzig noch die Höhe der Genugtuung.

[VERSION 1: Die Klägerin verlangt eine Genugtuung von Fr. 3'000'000.— (drei Millionen).]

[VERSION 2: Die Klägerin verlangt eine Genugtuung von Fr. 3'000'000.— (drei Millionen). Die Beklagte bietet Fr. 10'000.— (zehntausend) an.]

[VERSION 3: *Nichts, d.h. kein Antrag irgendeiner Partei.*]

Wie hoch ist Ihrer Meinung nach eine angemessene Genugtuung in diesem Fall, soweit Sie dies aufgrund der Angaben im Sachverhalt beurteilen können?

828 Frage 2: Beurteilung von Berufungsaussichten

Vergegenwärtigen Sie sich noch einmal den Sachverhalt der vorangehenden Frage (Unfall wegen Missachtung des Rechtsvortritts).

Wie sich im Strafverfahren herausstellt, wurde der fehlbare Lastwagenfahrer bereits vor sieben Jahren wegen Fahrens in angetrunkenem Zustand verurteilt. Zu einem Unfall kam es damals nicht. Der Lastwagenfahrer wurde damals mittels Strafbefehl zu einer bedingten Gefängnisstrafe von 14 Tagen verurteilt.

Soweit Sie dies beurteilen können, bereut der Lastwagenfahrer seine Tat aufrichtig. Eine Therapie wegen Alkoholsucht lehnt er ab, da es sich, wie er sagt, um zwei isolierte Einzelfälle handle und er kein Alkoholiker sei.

Das erstinstanzliche Gericht verurteilt den Lastwagenfahrer wegen fahrlässiger schwerer Körperverletzung, Fahrens in angetrunkenem Zustand und grober Verkehrsregelverletzung zu einer bedingten Gefängnisstrafe von vier Monaten.

[VERSION 1: Das Berufungsgericht bestätigt das Urteil des erstinstanzlichen Gerichts im Schuldpunkt. Im Strafpunkt reduziert es die Dauer der Gefängnisstrafe auf drei Monate, ordnet aber den unbedingten Vollzug an.]

[VERSION 2: Das Berufungsgericht bestätigt das Urteil des erstinstanzlichen Gerichts im Schuldpunkt. Im Strafpunkt senkt es die Dauer der Gefängnisstrafe auf drei Monate und gewährt den bedingten Vollzug.]

[VERSION 3: Das Berufungsgericht bestätigt das Urteil des erstinstanzlichen Gerichts in Schuld- und Strafpunkt vollumfänglich.]

Welcher Ausgang des Berufungsverfahrens war aufgrund des geschilderten Sachverhaltes am wahrscheinlichsten? (Es ist eine der angegebenen Möglichkeiten auszuwählen.)

- ☐ Das Berufungsgericht bestätigt das Urteil des erstinstanzlichen Gerichts im Schuldpunkt. Im Strafpunkt reduziert es die Dauer der Gefängnisstrafe auf drei Monate, ordnet aber den unbedingten Vollzug an.
- ☐ Das Berufungsgericht bestätigt das Urteil des erstinstanzlichen Gerichts im Schuldpunkt. Im Strafpunkt senkt es die Dauer der Gefängnisstrafe auf drei Monate und gewährt den bedingten Vollzug.
- ☐ Das Berufungsgericht bestätigt das Urteil des erstinstanzlichen Gerichts in Schuld- und Strafpunkt vollumfänglich.

829 Frage 3: Bewertung von statistischen Beweismitteln

Nachdem auch ohne Anlass Alkoholtests durchgeführt werden dürfen, lässt die Kantonspolizei Zürich alle Autofahrer, die eine Kontrollstelle am Stadtrand von Zürich passieren, ins Teströhrchen blasen. Erfahrungsgemäss ist einer von hundert getesteten Autofahrern betrunken.

Dieser Alcotest zeigt nicht ganz zuverlässig an, ob ein Autofahrer mehr als die zulässige Blutalkoholkonzentration aufweist. Es kommt vor, dass ein alkoholisierte Autofahrer nicht erwischt wird, weil sich das Röhrchen (leider) nicht verfärbt. Bedauerlicherweise verfärbt sich das Röhrchen aber auch gelegentlich bei nüchternen Autofahrern.

- Die Verfärbung tritt bei 95% aller Blaufahrer (zurecht) ein.

Besonderer Teil

- Die Verfärbung tritt aber auch (bedauerlicherweise) bei 5% aller nüchternen Fahrer ein.

Angenommen, einer von den vielen kontrollierten Autofahrern bläst ins Röhrchen, und dieses verfärbt sich.

Wie hoch ist die Wahrscheinlichkeit, dass die Polizei einen Blaufahrer erwischt hat?

- ☐ 0% - 25% Wahrscheinlichkeit, dass die Polizei einen Blaufahrer erwischt hat.
- ☐ 26% - 50% Wahrscheinlichkeit, dass die Polizei einen Blaufahrer erwischt hat.
- ☐ 51% - 75% Wahrscheinlichkeit, dass die Polizei einen Blaufahrer erwischt hat.
- ☐ 76% - 100% Wahrscheinlichkeit, dass die Polizei einen Blaufahrer erwischt hat.

830 Frage 4: Vergleichsangebot bewerten

Sie sind ReferentIn in einem Prozess zwischen zwei Baufirmen. Die Klägerin hat die Beklagte auf Zahlung eines Gewinnanteils von Fr. 150'000.— aus einem gemeinsamen Bauprojekt verklagt. Beide Baufirmen sind mittelgrosse Unternehmen mit einem Jahresumsatz von rund CHF 10 Mio. pro Jahr. Sie werden von fähigen Anwälten vertreten, die Ihnen nicht persönlich bekannt sind.

Wird der Fall bis zu einem Urteil durchprozessiert, so hat jede Partei mit eigenen Anwaltskosten von Fr. 30'000.— zu rechnen. Die Prozessentschädigung gemäss Gebührenverordnung beträgt rund Fr. 15'000.—; selbst bei einem vollständigen Obsiegen erhält die obsiegende Partei daher Fr. 15'000.— ihrer Anwaltskosten nicht ersetzt. Die von der unterliegenden Partei zu bezahlenden Gerichtskosten betragen Fr. 10'000. —.

Der Fall ist rechtlich einfach. Die Höhe des geschuldeten Betrages ist nicht umstritten, nur der Anspruch der Klägerin wird bestritten. Der Ausgang des Prozesses hängt einzig von der Aussage eines Schlüsselzeugen ab. Sie glauben, dass eine 50%ige Chance besteht, dass die Klägerin die vollen Fr. 150'000.— erhält und eine 50%ige Chance, dass die Klägerin Fr. 0.— erhält.

[VERSION 1: An der Vergleichsverhandlung bietet die Beklagte der Klägerin eine Zahlung von Fr. 50'000.— für den Rückzug der Klage an. Die Gerichtskosten würden von beiden Parteien je zur Hälfte übernommen und jede Partei trüge ihre eigenen Anwaltskosten. Für ihren Anteil an den Gerichtskosten und die eigenen Anwaltskosten muss jede Partei mit Fr. 10'000.— rechnen.

Glauben Sie, dass die Klägerin die Zahlung von Fr. 50'000.— akzeptieren sollte, um den Prozess zu erledigen? ☐ Ja ☐ Nein]

[VERSION 2: An der Vergleichsverhandlung bietet die Klägerin der Beklagten an, gegen eine Zahlung von Fr. 100'000.— die Klage zurückzuziehen. Die Gerichtskosten würden von beiden Parteien je zur Hälfte übernommen und jede Partei trüge ihre eigenen Anwaltskosten. Für ihren Anteil an den Gerichtskosten und die eigenen Anwaltskosten muss jede Partei mit Fr. 10'000.— rechnen.

Glauben Sie, dass die Beklagte Fr. 100'000. — zahlen sollte, um den Prozess zu erledigen?
☐ Ja ☐ Nein]

831 Frage 5: Berufsquote

Die Qualität eines Richters oder einer Richterin misst sich unter anderem am Anteil seiner/ihrer Entscheide, gegen die erfolgreich Berufung oder Rekurs erhoben wird (ob dieses Kriterium gerechtfertigt ist, bleibe dahingestellt).

Angenommen, es gäbe eine Rangliste aller schweizerischen Richter und Richterinnen. Zuoberst auf der Liste stünden die RichterInnen, deren Entscheide am wenigsten oft durch eine obere gerichtliche Instanz korrigiert werden.

Wo stünden Sie selber auf dieser Liste?

- ☐ Ich würde zu den besten 25% gehören (d.h. mehr als drei Viertel aller RichterInnen werden *öfter* durch eine höhere gerichtliche Instanz korrigiert als Sie).
- ☐ Ich würde zu den besseren 50% gehören (d.h. mehr als die Hälfte aller RichterInnen wird *öfter* durch eine höhere gerichtliche Instanz korrigiert als Sie).
- ☐ Ich würde zu den schlechteren 50% gehören (d.h. mehr als die Hälfte aller RichterInnen wird *seltener* durch eine höhere gerichtliche Instanz korrigiert als Sie).
- ☐ Ich würde zu den schlechtesten 25% gehören (d.h. mehr als drei Viertel aller RichterInnen werden *seltener* durch eine höhere gerichtliche Instanz korrigiert als Sie.)

Anhang C: Fragebogen 2004

832 Frage 1: Strafen und Massnahmen

X ist seit 1993 drei Mal unter anderem wegen strafbarer Handlungen gegen die sexuelle Integrität zu Gefängnisstrafen verurteilt worden (1993 zu 6 Monaten, 1997 zu 8 Monaten, 2000 zu 10 Monaten). Meist handelte es sich um Exhibitionismus, wobei er einmal auch ein kleines Mädchen ausgegriffen hatte.

Nach seiner Entlassung aus dem Strafvollzug Anfang November 2002 zog X bei einer ihm bekannten alleinerziehenden Mutter zweier Knaben (Jahrgang 1993 und 1996) als Untermieter ein und übernahm während der Abwesenheit der Mutter Betreuungsfunktionen. Während der Abwesenheit der Mutter zeigte X dem jüngeren Knaben mehrfach seinen erigierten Penis. Teilweise befriedigte sich X vor den Augen des jüngeren Knaben selber; einmal beobachtete der ältere Knabe dies. X versuchte, den jüngeren Knaben zum Oralsex zu zwingen, sah aber davon ab, als dieser sich wehrte.

X schrieb 2003 in einem Primarschulhaus an fünf Orten seine Telefonnummer auf, mit einem pornographischen Text und der Aufforderung, "ruft mich an, für Mädchen von 7 - 19 Jahren".

Welche der zur Auswahl stehenden Strafen/Massnahmen betrachten Sie in diesem Fall als angemessen? [VERSION 1: Die lebenslängliche Verwahrung ohne Möglichkeit der vorzeitigen Entlassung steht nicht zur Verfügung.]

- ☐ Drei Jahre Gefängnis (unbedingt) mit begleitender ambulanter Psychotherapie.
- ☐ Drei Jahre Gefängnis, wobei der Vollzug der Strafe zugunsten einer Verwahrung auf unbestimmte Zeit aufgeschoben wird. Der Angeklagte wird entlassen, wenn feststeht, dass er keine Gefahr für die Gesellschaft mehr darstellt.
- ☐ [VERSION 2: Drei Jahre Gefängnis, wobei der Vollzug der Strafe zugunsten einer lebenslänglichen Verwahrung ohne Möglichkeit der vorzeitigen Entlassung aufgeschoben wird.]

833 Frage 2: Haftung

Bei der Erstellung von Mietshäusern liess der verantwortliche Bauleiter B. Aushubmaterial in einem Tälchen ablagern. Das Aushubmaterial staute den durch das Tälchen fliessenden Bach. Der flache Teich, den der Bach vorher gebildet hatte, wurde dadurch viel grösser und tiefer (die Tiefe vergrösserte sich von 25 cm auf 4 m).

[VERSION 2: Im Februar, als der Teich teilweise zugefroren war, sahen die 7 bzw. 11 Jahre alten Brüder Luca und Marco einen verlassenen Ball auf dem Eis liegen. Der jüngere Knabe betrat das Eis, um den Ball zu holen. Dabei brach er im Eis ein und ertrank zusammen mit seinem Bruder, der ihm zu Hilfe kommen wollte.]

Handelt der Bauleiter fahrlässig, wenn er den Teich weder durch eine Warntafel noch durch eine Abschränkung sichert?

- ☐ Ja.
- ☐ Nein.

[Je 50 % jeder Variante der Frage wurde von einem der beiden untenstehenden Bilder, das mit „Bauleiter B.“ bezeichnet war, versehen.]



Bauleiter B.

834 Frage 3: Strafzumessung

Das Ehepaar E. ist seit vielen Jahren verheiratet. Die Ehe ist nicht mehr glücklich und der Ehemann hat eine Geliebte. Er würde sich am liebsten scheiden lassen, verzichtet aber aus finanziellen Gründen darauf.

Die Ehefrau leidet an einer Erdnuss-Allergie. Diese kann beim Verzehr von Erdnüssen zu einem allergischen Schock und zum Tod führen. Ihr Ehemann weiss dies.

Das Stammlokal des Ehemannes macht eine hausgemachte Salatsauce („House-Dressing“), die einen geringen Anteil an pürierten Erdnüssen enthält. Der Ehemann weiss von den Erdnüssen im House-Dressing.

Auf Wunsch der Ehefrau geht das Ehepaar gemeinsam in das Stammlokal des Ehemannes.

Als die Ehefrau einen Salat bestellt, fragt der Kellner, ob sie italienische oder französische Sauce oder das House-Dressing möchte. Die Ehefrau weiss nicht, dass das House-Dressing Spuren von Erdnüssen enthält und erkundigt sich auch nicht danach.

[VERSION 1: Bevor sich die Frau entscheiden kann, empfiehlt ihr Mann das House-Dressing. Die Frau entscheidet sich daraufhin für das House-Dressing. Sie erleidet einen allergischen Schock und stirbt. Hätte der Mann nichts gesagt, so hätte seine Frau die italienische Sauce bestellt.]

[VERSION 2: Bevor sich die Frau entscheiden kann, empfiehlt ihr Mann das House-Dressing. Die Frau entscheidet sich dann auch für das House-Dressing. Sie erleidet eine allergische Reaktion, die allerdings sehr schwach ausfällt und ausser einem kurzzeitigen Unwohlsein keine weiteren Folgen hat. Hätte der Mann nichts gesagt, so hätte seine Frau die italienische Sauce bestellt.]

E. Der prozentuale Anteil Kombis an der Gesamtheit in der Schweiz zugelassener schwarzer Personenwagen der Marke Opel.]

Welche der drei zusätzlichen Informationen (C, D oder E) hilft Ihnen entscheiden, ob die Fussgängerin einen Opel oder einen Ford gesehen hat? Sie können nur eine einzige zusätzliche Information erfahren. Bitte kreuzen Sie ihre Wahl an.

836 Frage 5: Verschulden

Während die Automobilistin A. sich auf einer dicht befahrenen Strasse innerorts befindet, bedient sie das Autoradio, um einen neuen Sender einzustellen.

Sie übersieht deshalb, dass das Auto vor ihr angehalten hat, um dem Fussgänger F. das Überqueren der Strasse auf dem Fussgängerstreifen zu erlauben.

Der von der Automobilistin A. gelenkte Wagen prallt in das Heck des vor ihr stehenden Autos. Dieses wird durch die Wucht des Aufpralls nach vorne geschoben und verletzt den Fußgänger F. leicht. Die beiden Fahrzeuglenker bleiben unverletzt.

Wie beurteilen Sie auf einer Skala von 1 bis 7 das Verschulden der Automobilistin A.?

geringes 1 ----- 2 ----- 3 ----- 4 ----- 5 ----- 6 ----- 7 schweres
Verschulden Verschulden

Was ist Ihrer Ansicht nach eine angemessene Strafe für (die nicht vorbestrafte) Automobilistin A.?

□ Busse in der Höhe von Fr.

☐ Bedingte Freiheitsstrafe von ____ Tagen.

☐ Busse in der Höhe von Fr. _____ und bedingte Freiheitsstrafe von _____ Tagen.

(Über administrative Massnahmen, insbesondere Führerausweisentzug, können Sie nicht entscheiden.)

[Je 50 % der Fragebogen waren mit einem der beiden untenstehenden Bildern versehen.]



Automobilistin A.

Besonderer Teil

Lebenslauf

1986-1992	Realgymnasium Rämibühl, Zürich (Matura Typus B)
1993-1998	Universität Zürich, Rechtswissenschaftliche Fakultät (Lizentiat)
1999-2001	Bezirksgericht Horgen, Zürich (Auditor und juristischer Sekretär)
2001	Zulassung als Anwalt in Zürich
2001-2002	University of Michigan Law School, Ann Arbor, USA (LL.M.)
2002-2004	Meyer Lustenberger Rechtsanwälte, Zollikon, Zürich
2005-	Delegierter des Internationalen Komitees des Roten Kreuzes